

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 08:58:20 ; Search time 5973 Seconds
(without alignments)
11178.853 Million cell updates/sec

Title: US-10-070-386-1

Perfect score: 1378

Sequence: 1 gtcgacgtgggtgctgcatat.....accatcaatcaactaaca 1378

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1378	100.0	1378	6	BD013073
2	48.2	3.5	2000	6	AX655393
3	42.2	3.1	7218	6	166494
4	41.8	3.0	231585	2	AC103186
5	41.8	3.0	329549	2	AC117331
6	41.6	3.0	164377	2	AC136809
7	41.6	3.0	226170	2	AC105687
8	40	2.9	6197	6	AX344849
9	39.8	2.9	221649	2	AC137392
10	39.8	2.9	262854	2	AC103016
11	39.8	2.9	342084	2	AC128752
12	39.6	2.9	81624	3	AC099763
13	39.6	2.9	290066	2	AC129421
14	39.4	2.9	81921	10	AL805938
15	39.4	2.9	186858	10	AC130210
16	39.4	2.9	195411	10	AL669958
17	39.4	2.9	218807	10	BX511249
18	39	2.8	237119	10	AC115291
19	38.8	2.8	16512	3	CEU49795

C 20	38.8	2.8	34919	3	AF100663
C 21	38.8	2.8	139549	2	AC114644
C 22	38.6	2.8	129489	10	AL713860
C 23	38.6	2.8	143148	10	AL607034
C 24	38.6	2.8	201295	2	BX901879
C 25	38.6	2.8	225552	2	BX897745
C 26	38.6	2.8	233819	2	AC095098
C 27	38.6	2.8	234873	2	AC137028
C 28	38.6	2.8	240425	10	CNS07XOT
C 29	38.6	2.8	272563	2	AC095469
C 30	38.4	2.8	195611	10	AC121971
C 31	38.4	2.8	334371	2	AC109571
C 32	38.2	2.8	106259	2	AC151909
C 33	38.2	2.8	144298	10	AC127682
C 34	38.2	2.8	180907	9	AC131157
C 35	38.2	2.8	202270	10	AC102241
C 36	38.2	2.8	253768	2	AC125994
C 37	38.2	2.8	302921	2	AC090998
C 38	38	2.8	194370	2	AL645472
C 39	38	2.8	212543	5	BX663615
C 40	38	2.8	220638	10	AL645723
C 41	38	2.8	257760	2	AC126638
C 42	37.8	2.7	114999	9	AP003421
C 43	37.8	2.7	142148	9	AC025465
C 44	37.8	2.7	167912	9	AC060232
C 45	37.8	2.7	170066	9	AL161730

ALIGNMENTS

RESULT 1
BD013073
LOCUS
DEFINITION
Regulatory sequences and expression system functional in mold fungi.

ACCESSION
BD013073.1 GI:22093262
VERSION
WO 0118219-A/1.
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.

REFERENCE
1 (bases 1 to 1378)
Watanabe, M. and Murakami, T.
Regulatory sequences and expression system functional in mold fungi
Patent: WO 0118219-A 1 15-MAR-2001;
JOURNAL
MEIDI SEIKA KAISHA LTD./MANABU WATANABE, TAKESHI MURAKAMI
COMMENT
OS Mycelia sterilia
PN WO 0118219-A/1
PD 15-MAR-2001
PF 07-SEP-2000 WO 2000JP006104
PR 07-SEP-1999 JP 99P 252851
PI MANABU WATANABE, TAKESHI MURAKAMI
PC C12N15/80, C12N1/15, C12P21/00//((C12N15/80, C12R1:645),
PC C12R1:645),
PC (C12P21/00, C12R1:645)
CC
FH

FEATURES
source
Key Location/Qualifiers.
1..1378
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1378; DB 6; Length 1378;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGTGGGTGATATCATGCTGTGTCCAAACTGTAGTACTACGAAT 60
DB 1 GTCGACGTGGGTGATATCATGCTGTGTCCAAACTGTAGTACTACGAAT 60

61 GAGGAAAGAAACGGTGGTGTGTGGCAGCTGAAAGACTGAAAGAGGACCAAGATTAATTC 120
61 GAGGAAAGAAACGGTGGTGTGTGGCAGCTGAAAGACTGAAAGAGGACCAAGATTAATTC 120
121 ACAAATGCGATAGAGGTTCATCAATGCTTTTCAAGAGACAGTTGCATCTACCTGGTGT 180
121 ACAAATGCGATAGAGGTTCATCAATGCTTTTCAAGAGACAGTTGCATCTACCTGGTGT 180
181 TCCCTCTCTGTTGTACAGATCAAGTATCGGATGACACCCACCCCGACGAAATCTCTG 240
181 TCCCTCTCTGTTGTACAGATCAAGTATCGGATGACACCCACCCCGACGAAATCTCTG 240
241 GAGTTCAAAAGAGGGTGTCTCTACCGGCAATTAAGTATAGATGAGGATGAGGTTTGA 300
241 GAGTTCAAAAGAGGGTGTCTCTACCGGCAATTAAGTATAGATGAGGATGAGGTTTGA 300
301 GGTGAAAGGTGATTAACAGACATGAGACAAAGAAAATAACAGGTTGTATGCTTCCCGT 360
301 GGTGAAAGGTGATTAACAGACATGAGACAAAGAAAATAACAGGTTGTATGCTTCCCGT 360
361 GCTTAATAAGTATGATTCAGAGACACACAGCCGAAAGAACCGATGCTGTCTGAGGG 420
361 GCTTAATAAGTATGATTCAGAGACACACAGCCGAAAGAACCGATGCTGTCTGAGGG 420
421 TTCTTTAGAGTCTACATGATGATAACGGTATGATGATAAATCAATCAATGAGGATCA 480
421 TTCTTTAGAGTCTACATGATGATAACGGTATGATGATAAATCAATCAATGAGGATCA 480
481 AATTATTCCTGACGCTACATCGCTTTCTTCCGATCTTGTCTTAAATATATGTCCTGTC 540
481 AATTATTCCTGACGCTACATCGCTTTCTTCCGATCTTGTCTTAAATATATGTCCTGTC 540
541 GAAGTGTGCGTATGCTTCTGATCAATGCTTCTTCCGATGAGTCTTGAAGTCTTGAAG 600
541 GAAGTGTGCGTATGCTTCTGATCAATGCTTCTTCCGATGAGTCTTGAAGTCTTGAAG 600
601 CGTTTGTAGACCTACATGATGATGACATCTTAAAGAGGAGATGAGACATTTTCTAAG 660
601 CGTTTGTAGACCTACATGATGATGACATCTTAAAGAGGAGATGAGACATTTTCTAAG 660
661 ATCCATATAGGCAATTTGGGCGCTTAAGTCCGATTTGAAGAGATTAAGGGGGTGTGAAG 720
661 ATCCATATAGGCAATTTGGGCGCTTAAGTCCGATTTGAAGAGATTAAGGGGGTGTGAAG 720
721 GTGTGCAAAAGAGGCTGATTTGGCTATATACAGCCGCTTAAGAGAGTGTGGAAGTGT 780
721 GTGTGCAAAAGAGGCTGATTTGGCTATATACAGCCGCTTAAGAGAGTGTGGAAGTGT 780
781 CTGACAGCTGTGAATTAACGTCATCTGCTTGAATATGTCACCTAATGTCAAGCAGATCA 840
781 CTGACAGCTGTGAATTAACGTCATCTGCTTGAATATGTCACCTAATGTCAAGCAGATCA 840
841 TGTGATTTGGGTTAAATGAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 900
841 TGTGATTTGGGTTAAATGAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 900
901 AAAAGGAAGCTGAAGCTGAAGCTGATCAAGAAATTAAGCTGTTGAATATCAAGCTTTA 960
901 AAAAGGAAGCTGAAGCTGAAGCTGATCAAGAAATTAAGCTGTTGAATATCAAGCTTTA 960
961 CCCAATTCAGTGTCTCAAGAGGTGTCTGATATGCTGAGAGCTTCCCTGTGCGATTTG 1020
961 CCCAATTCAGTGTCTCAAGAGGTGTCTGATATGCTGAGAGCTTCCCTGTGCGATTTG 1020
1021 AACTATTTTCAATGATGAGGAGCAATGCAATCTTATTTTCAATTTGAATCTAAATCTT 1080
1021 AACTATTTTCAATGATGAGGAGCAATGCAATCTTATTTTCAATTTGAATCTAAATCTT 1080
1081 GTAGAGGTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081 GTAGAGGTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141 ATACAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200

1141 ATACAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1201 CAAGACATGAGGCGTACGTAGATACATAGATGATGATGATGATGATGATGATGATGAT 1260
1201 CAAGACATGAGGCGTACGTAGATACATAGATGATGATGATGATGATGATGATGATGAT 1260
1261 CATTAATTCCTAAGAGGTTCTCTTCTTCTGATGATGATGATGATGATGATGATGATGAT 1320
1261 CATTAATTCCTAAGAGGTTCTCTTCTTCTGATGATGATGATGATGATGATGATGATGAT 1320
1321 CAGGAATCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1378
1321 CAGGAATCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1378

RESULT 2

AX655393 LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 5263 from Patent WO0300898.
ACCESSION AX655393
VERSION AX655393.1 GI:29158207
KEYWORDS

SOURCE

ORGANISM Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
Katesgiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES location/Qualifiers

source

1..2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Query Match 3.5%; Score 48.2; DB 6; Length 2000;
Best Local Similarity 10.0%; Pred.No.0.0074;
Matches 79; Conservative 339; Mismatches 372; Indels 1; Gaps 1;
585 CTTAGACAGAGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
28 SCHARMSHMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSK 87
645 AGACATTTTCTAAGGATCATAATGAGCATTTGGCGCTTAAGTCCGATTTGAAGAGATTA 704
88 KMKKRRYKRSKMRGRGRMRBSRMRBMGRYRRCARSGRMAGSGSRMMSKMSKMSKMSK 147
705 GGGGGGTGGAAGTGTGTGTCGCAAGAGAGTGCATGATGATGATGATGATGATGATGATGAT 764
148 CVARGGSCRRKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSK 207
765 GTGGGCTAGACAGTGTCTGACGTGATGATGATGATGATGATGATGATGATGATGATGAT 824
208 RSATSYRYTSSKRYTGYKMTYSASRCMBATYTTISYASCSYTTMCRRSRMRMRMRMR 267
825 TGTACAGACAGTCAAAATGCTGATTTGGTTAAATGAGCATGATGATGATGATGATGATGAT 884
268 KMRMSRSYGMWSYGMWSYGMWSYGMWSYGMWSYGMWSYGMWSYGMWSYGMWSYGMWSYGM 327
885 ACCTTTAGATCTAATGATTTAAAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 944
328 KMYRYGYKMRKMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMRGMR 387
945 AATTACAAGTGTATTAACCAATTCAGTGTGCAAGAGGATGATGATGATGATGATGATGATGAT 1004
388 RKMCMKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYKRRYK 447

[illegible]

Oy	907	AAGCTGAACGTAACCTGTGAGAATAATAGCCTTTGGAAATACAACGTTGATAACCAAT	966
Dd	1150	RRR	1091
Oy	967	TCACTGCAGAGGCGTCCATATNGCTGGAGCTTCCTGTGCATTGTGGGTAACTAT	1026
Dd	1090	RRR	1031
Oy	1027	TTCATAGTGCGGCAGAAATGCAACTCTATTTTCAATGATGTA	1069
Dd	1030	TTAATTCGTGAGCGTATGCCAAGAACGAAAATAATGTTATA	988
RESULT 4			
LOCUS	AC103186		
DEFINITION	Rattus norvegicus clone CH230-107D17, *** SEQUENCING IN PROGRESS		
ACCESSION	AC103186		
VERSION	AC103186.5 GI:30578507		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus Eukaryota Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 231585)		
REFERENCE	Murphy J, Marle H, Metzker M, Lee S, Abramson S, Adams C, Alder J, Allen C, Allen H, Meisner M, Lee S, Amin A, Angiano D, Anyaladebe V, Ayoyagi A, Ayodeji I, Bacc E, Baden H, Baldwin K, Balaramanikam D, Barber M, Barnstead M, Benahmed F, Bialwalo K, Blair D, Blankenburg K, Bluth P, Brown M, Byrant N, Buha Y, Burck C, Burck P, Butrell K, Calderon E, Cardenas V, Carter K, Cavazos I, Caesar H, Centler A, Chacko J, Chavez D, Chen G, Chen R, Chen X, Chen Z, Chu J, Cleaveland C, Cockrell R, Cox C, Coyle M, Cree A, D Souza L, David M, Davis C, Davy-Carroll L, De Anda C, Dedrich D, Delgado O, Denison S, Deramo C, Ding Y, Dinh H, Divya K, Diexper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escoto M, Eugene C, Evans C.A., Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Frazer C.M., Gadisi A, Gante R, Garcia A, Garner F, Garza M, Georgescu E, Geor K, Gill R, Grady W, Guerra M, Guevara W, Gunaratne P, Haaland W, Hamill C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hudson S.L., Hodgson A, Hogues M, Hollins B, Howells S, Huliy S, Hume J, Idlebird D, Jackson A, Jackson S, Jacob L, Jiang H, Johnson B, Johnson R, Jolyet A, Karpachy S, Kelly S, Kelley S, Khan Z, King L, Kovac K, Kwis C, Kraft C.L., Lebow H, Levay J, Lewis L, Li Z, Liu J, Liu J, Liu Y, Liu Y, London P, Longacre S, Lopez J, Lorenshewa L, Loulsged H, Lozada R.J., Lu X, Ma J, Maheshwari M, Mahindratne M, Mahmood M, Malloy K, Mangum A, Mangun B, Mapua P, Martin K, Martin R, Martinec E, Mahoney S, McLeod M.P., McNeill T.Z., Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Neir L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nowokolenko O, Okwuonu G, Olarinpoju A, Pal S, Parks K, Patel N, Paul H, Perez A, Perez L, Plannkocch C, Platopier F, Polndexter A, Popovic D, Primus E, Pu L.-L., Puzio M, Quiroz J, Rachin E, Reeves K, Regier M.A., Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs F, Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S.J., Sanders W, Savery G, Scherer S, Scott C, Shatsman S, Shen H, Shetty J, Shivatsbeyn A, Sisson I, Sitter C.D., Smaj S,D., Sneed A, Sodergren E, Song X.-Z., Soelle R, Sosa J, Steidle M, Strong R, Sutton A, Swalek A, Taber P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Uemami K, Valas R, Vera V, Villaena D, Waldron L, Walker B, Wang J, Wang O, Wang S, Warren J, Warren R, Wei X, White F, Williams G, Willson R, Wlezky R, Wooden H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V,		

```

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 231585)
Direct Submission
Worley,K.C.
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 231585)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23365296.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJRW
Center clone name: CH230-107D17
----- Summary Statistics
Assembly program: ATras 3.0/
Consensus quality: 220133 bases at least Q40
Consensus quality: 223263 bases at least Q30
Consensus quality: 224520 bases at least Q20
Estimated insert size: 239728; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 230369: contig of 230369 bp in length
* 230370 230469: gap of unknown length
* 230470 231585: contig of 1116 bp in length.
Location/Qualifiers
1..231585
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-107D17"
1..2354
/note="wgs_end_extension
clone_end:Sp6"
5337..6318
/note="clone boundary
clone_end:Sp6
site:BCOR1
end_sequence:BH359310"
FEATURES
source
misc_feature
misc_feature
misc_feature

```

```

misc_feature      complement(1221205..227963)
                  /note="clone_boundary
                  clone_end:17
                  site:ECORI
                  end_sequence:BH359309"

ORIGIN
Query Match      3.0%; Score 41.8; DB 2; Length 23155;
Best Local Similarity 47.2%; Pred. No. 0.79;
Matches 127; Conservativity 0; Mismatches 142; Indels 0; Gaps 0;

Oy    1110 CACTTACGACATCATGTGGGGCTCAACAACGTTATCAGCTTCATATGAGAGTGCGGCATTGA 1169
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     168436 CACACACACACACACACACACACACACACACACAGAGAGAGAGAGAGAGAGAGAGA 168495
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy    1170 AGTAGCTACCCGATCGAACCCGGAAGCGGTTCAAGACATGGGCCGTACGTTAGATACATAGA 1229
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     168496 CAGAGAGACAGAGAGACAGAGAGACAGAGAGACAGAGAGACAGACAGAGAGACAGAGA 168555
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy    1230 GTCATAGAAACATTAAGAGAGCTTGTAAGACCATTCAAATCCTTAAGGCTCTCTCTTT 1289
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     168556 GGCAAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 168615
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy    1250 CTGCATCATCATCAAGATCATCACTCACTCAACACGAGAACCTTTCTATCTTCCCTATAGCA 1349
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     168616 CTGAGGAACATTTTAAAAAAGCTCATCACAGTAAATAATTAACATCCAGAAATCTATA 168675
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy    1350 ATTCCCAAAACCCATCAATCAACCTTACGA 1378
Db     168676 CTTTTCATTAATACTTCACTGACTTAAAA 168704
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
AC117331      329549 bp   DNA          linear   HTG 09-OCT-2002
LOCUS        Rattus norvegicus clone CH230-242a22, *** SEQUENCING IN PROGRESS
DEFINITION   *** 4 unordered pieces.
ACCESSION   AC117331
VERSION     AC117331.4 GI:23194666
KEYWORDS    HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 329549)
REFERENCE   1 (bases 1 to 329549)
AUTHORS    Munzy,D.,Marie,,Metzker,M.Lee,,Abramson,S.,Adams,C.,Alder,J.,
            Allen,C.,Allen,H.,Alsbrooks,S.,Amir,A.,Anguiano,D.,
            Anyalebechi,V.,Ayogisi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
            Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
            Bisswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
            Bryant,N.,Buhaq,C.,Burck,P.,Butrell,K.,Calderon,E.,
            Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,
            Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
            Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cre,A.,D'Souza,L.,
            Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Andra,C.,Dederich,D.,
            Delgado,O.,Denison,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,
            Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duval,B.,Eaves,K.,
            Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
            Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
            Frazer,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garner,T.,Garza,M.,
            Gebregergis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,
            Gunaratne,P.,Haaland,W.,Hamill,C.,Hamilton,C.,Hamilton,K.,
            Harvey,Y.,Haylak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
            Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
            Hollins,B.,Howells,S.,Huik,S.,Hume,J.,Idlebird,D.,Jackson,A.,
            Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolive,A.,
            Karpathy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Koyr,C.,
            Kows,C.,Kraft,C.L.,Lebow,H.,Leyan,J.,Lewis,L.,Li,Z.,Liu,J.,
            Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,U.,
            Lorenshewka,L.,Louisegh,H.,Lozada,R.J.,Lu,X.,Ma,J.,
            Maheshwari,M.,Mahindaratne,M.,Mahmoud,M.,Mallory,K.,Mangum,A.,
            Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,

```


Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinth, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Fotes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregziabher, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, B., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idebrld, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpethy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowis, C., Kraft, C.L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresunhwa, L., Loulseged, H., Lozadó, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangus, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mlawnslevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Mloaslevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nsoekiemeh, O., Okwunodu, G., Olarunpasegon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldexter, A., Popovic, D., Primus, E., Pu, L., L., Piazio, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scheerer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Silter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Szatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Velas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, D., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weise, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 164377)
Rat Genome Sequencing Consortium.
Submitted (08-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 164377)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:24796682.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCSF

RESULT 7
AC105687/c
LOCUS AC105687 226170 bp DNA linear HTG 08-OCT-2002
DEFINITION Rattus norvegicus clone CH230-1111, *** SEQUENCING IN PROGRESS ***
3 unordered pieces.
AC105687
AC105687.3 GI:23101526
VERSION HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 226170)
Munzy,D,Marie., Metzger,M,lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Bismalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Butrell,K., Caesar,H., Canter,A., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabiela,A., Ganta,R., Garcia,A., Garner,T., Garza,W., Gebregorjios,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M., Hollins,B., Howell,S., Huijck,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshew,L., Loulsegad,H., Lozano,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,M., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Mijic,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelmehe,O., Okwunonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C., Plopper,F., Polindexter,A., Popovic,D., Pritts,E., Pu,L., P., Puato,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steadler,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,D., Tingey,A., Trejos,Z., Umamang,K., Valsar,R., Vera,V., Villasana,S., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K., Yu,F., Zhang,Y., Zhou,J., Zhou,X., Zhou,S., Yen,J., Yoon,L., Yoon,V., Niedermauern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.
TITLE Unpublished
JOURNAL Direct Submission
AUTHORS 2 (bases 1 to 226170)
Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
3 (bases 1 to 226170)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 18, 2002 this sequence version replaced gi:21743897.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNNC
Center clone name: CH230-1111
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 198186 bases at least Q40
Consensus quality: 202314 bases at least Q30
Consensus quality: 204704 bases at least Q20
Estimated insert size: 229286; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 7584: contig of 7584 bp in length
* 7585 7684: gap of unknown length
* 7685 224621: contig of 216937 bp in length
* 224622 224721: gap of unknown length
* 224722 226170: contig of 1449 bp in length.
FEATURES
source
1..226170
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1111"
/note="wgs_contig"
misc_feature
45780..46902
ORIGIN
Query Match 3.0%; Score 41.6; DB 2; Length 226170;
Best Local Similarity 49.1%; Pred. No. 0.91;
Matches 110; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 1155 AGAGTGGCGGCAATGAAGTACCTACCGCATGACCCGGAAGCGGTTCAAGACATGGGCGT 1214
61562 AGATGCGCTGGAATTAAGAGATGACAAAGAGGAGGAGACATATGAAGATGATCAGG 61503
QY 1215 ACCTAGATACATAGACTCATAGAAACATATAAGGAGACTTGAAGAACATTCATCAATCTTA 1274
61502 AACTGTATGACACATATTAAGTACGATTCATGAAGAACATCATTTCTATTAAATATATTA 61443
QY 1275 GGGTCTCTCTTCTTCTGATCATCAAGAAATCATACATCAACACAGAACTTCTTCT 1334

```

Db      61442  CGTATCATGTTTCTTCTAATAAATAAAGTAAAGATACAGAGAACTCATGA 61383
Qy      1335  ATCTTCCTATAGCAATCCCAAAACCATCATCACTAACA 1378
Db      61382  GACTGAACTAGAGAAATCCCAACCTCTAGATACATGACA 61339

RESULT 8
AX344849/c
LOCUS   AX344849 6197 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 274 from Patent WO0200927.
ACCESSION AX344849
VERSION   AX344849.1 GI:18492735
KEYWORDS
SOURCE    synthetic construct
           synthetic construct
           other sequences; artificial sequences.
ORGANISM  1
REFERENCE 1
AUTHORS   Olek, A., Piepenbrock, C. and Berlin, K.
TITLE     Diagnosis of diseases associated with development genes
JOURNAL   Patent: WO 0200927-A 274 03-JAN-2002;
           Epigenomics AG (DE)
FEATURES
           location/Qualifiers
           1..6197
           /organism="synthetic construct"
           /mol_type="unassigned DNA"
           /db_xref="taxon:32630"
           /note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match 2.9%; Score 40; DB 6; Length 6197;
Beet Local Similarity 51.7%; Pred. No. 2.5;
Matches 91; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy      1200  TCAGACATGGCGTACGTAGATACATGAGTATGAAACATAAAGAGCTTGAGAA 1259
Db      3228  TAAATTAATTAACAACAACAACAACAACAACAACAACAATAATTAATTAACA 3169
Qy      1260  CCATTCATATCCTAAGAGGTCTCTCTCTTCTTGCAATCAATCAAGATCATCAGCA 1319
Db      3168  CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3109
Qy      1320  CCAGAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1375
Db      3108  CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3053

RESULT 9
AC137392
LOCUS   AC137392 221649 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
           ***, 5 unordered pieces.
ACCESSION AC137392
VERSION   AC137392.1 GI:25138502
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE    Rattus norvegicus (Norway rat)
           Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE 1 (bases 1 to 221649)
AUTHORS   Muzny, D., Marle, J., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
           Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
           Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
           Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F.,
           Bissulo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
           Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
           Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
           Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
           Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L.,
           Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D.,
           Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

```

```

Drepper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garra, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensushewa, L., Louissege, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNell, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundiada, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackelamah, O., Okwou, G., Olarnpungoon, A., Pal, S., Parks, K.,
Paeternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L.,
Puafo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Thor, P., Taylor, C.,
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wiczysk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 221649)
Rat Genome Sequencing Consortium.
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KZNG
Center clone name: CH230-unknown
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 150999 bases at least Q40
Consensus quality: 156151 bases at least Q30
Consensus quality: 159735 bases at least Q20
Estimated insert size: 155933; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

```

```

*****
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 155377: contig of 155377 bp in length
* 155378 155477: gap of unknown length
* 155478 217856: contig of 62379 bp in length
* 217857 217956: gap of unknown length
* 217957 218983: contig of 1027 bp in length
* 218984 219083: gap of unknown length
* 219084 220305: contig of 1222 bp in length
* 220306 221649: gap of unknown length
* 221649 220406: contig of 1244 bp in length.
*
* Location/Qualifiers
*
*   source          1..221649
*                   /organism="Rattus norvegicus"
*                   /mol_type="genomic DNA"
*                   /db_xref="taxon:10116"
*                   /clone="CH230-unknown"
*                   1..2920
*                   /note="wgs contig"
*                   32902..34790
*                   /note="wgs contig"
*                   69822..72586
*                   /note="wgs contig"
*                   89976..91515
*                   /note="wgs contig"
*                   155478..156627
*                   /note="wgs contig"
*                   172910..174635
*                   /note="wgs contig"
*                   197144..200776
*                   /note="wgs contig"
*
* ORIGIN
*
* Query Match      2.9%; Score 39.8; DB 2; Length 221649;
* Best Local Similarity 50.8%; Pred. No. 3.2;
* Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
*
* Oy 11 GTGGTATATCATGCGTGTGTGTCCAAACCTGTGTAAGTACGATGAGAAAGAA 70
*      |||
* Db 36230 GTGTTTGATATCTTCTCTTGTAGATATGTGTTAAGAGAACATGAAAAAGTCA 36289
*
* Oy 71 ACGGTGTGTGTCGACGCTGAAAGACTGAAAGAGAGCCAAATATTTACATGCGAT 130
*      |||
* Db 36290 AAGGCTTAAGCAAGGCTTTCACAAAGAAAGAACTTTCAACAGAAAGAGAAAGCTGA 36349
*
* Oy 131 ACGGTGATCATGCTGTGTTCAAGAGAGACGTTGATCTACCTGGTTCCTCTTC 190
*      |||
* Db 36350 AAGCTGATCAAAACTGTTCTCCACCTCCCTTGTTTCATCTCTCTCTCTCTCT 36409
*
* Oy 191 GTTGTAC 197
*      |||
* Db 36410 GTTGAC 36416
*
* RESULT 10
* AC103016 262254 bp DNA linear HTG 13-MAY-2003
* LOCUS AC103016
* DEFINITION Rattus norvegicus clone CH230-1711L9, *** SEQUENCING IN PROGRESS
* *** 5 unordered pieces.
* ACCESSION AC103016
* VERSION AC103016.8 GI:30580739
* KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
* SOURCE Rattus norvegicus (Norway rat)
* ORGANISM Rattus norvegicus

```

```

REFERENCE
AUTHORS
1 (bases 1 to 262254)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
2 (bases 1 to 262254)
Muzny,D,Marrie,M,et alker,M, Lee, A, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flaggs, N., Forbes, L., Foster, T., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Loulleged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwuonu, G., Olamugbo, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sisson, L., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sores, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, E., Umami, K.,
Valae, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, D.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Woodden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhuser, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G., and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 262254)
Morley, K.C.
Direct Submission
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 262254)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:24819243.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,

```


AUTHORS	Morley, K. C.
TITLE	Direct Submission
JOURNAL	Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 342084)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Nov 15, 2002 this sequence version replaced gi:23803473. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
	----- Genome Center
	Center: Baylor College of Medicine
	Center code: BCM
	Web site: http://www.hgsc.bcm.tmc.edu/
	Contact: hgsc-help@bcm.tmc.edu
	----- Project Information
	Center project name: KALG
	Center clone name: CH230-466L2
	----- Summary Statistics
	Assembly program: Phrap; version 0.990329
	Consensus quality: 143887 bases at least Q40
	Consensus quality: 144547 bases at least Q30
	Consensus quality: 144951 bases at least Q20
	Estimated insert size: 177730; sum-of-contigs estimation
	Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

	* NOTE: Estimated insert size may differ from sequence length
	* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
	* NOTE: This is a 'working draft' sequence. It currently
	* consists of 1 contigs. Gaps between the contigs
	* are represented as runs of N. The order of the pieces
	* is believed to be correct as given, however the sizes
	* of the gaps between them are based on estimates that have
	* provided by the submitter.
	* This sequence will be replaced
	* by the finished sequence as soon as it is available and
	* the accession number will be preserved.
	1 342084: contig of 342084 bp in length.
FEATURES	
Source	Location/Qualifiers
	1..342084
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-466L2"
	184048..186150
	/note="wgs_contig"
ORIGIN	
Query Match	2.9%; Score 39.8; DB 2; Length 342084;
Best Local Similarity	50.8%; Pred. No. 3.3;
Matches	95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
0y	11 GTGATATATCATGAGCTGTGTGTCACAAACTGTGTAGTGAATGCAATGAGAAAGAA 70
Db	90802 GTGTTGGATCTCTTCCTTTGTAGAAATATGTTAAAGAGACATGAAAAAGTCAGAG 90861
0y	71 ACGGTGCTGTGTGTGACGTGAAGAATCGAAGAGAGACCAAAATATTAATTCACATCGAT 130
Db	90862 AAGCTTAGACAAAGCTTTACAAAGAGAAAGCTTTCAACAAGAAAGAGAAAAAGCTGA 90921

QY	131	ACGGTGTGCATCAATGCTTGTTCGAAGAGACGCGTGCATGCTGCTGTTC	130
Db	90922	AAGCTGAAATCAAACTGTTCTCCACCTCCCTGTTGTCATCTCTCTCTCT	9098
QY	191	GTGTGTAC 197	
Db	90982	GTTCGAC 90988	
RESULT 12			
AC099763/c	AC099763	81624 bp DNA	linear INV 20-NOV-2001
LOCUS	Caenorhabditis briggsae	cosmid CB022010,	complete sequence.
DEFINITION	AC099763		
ACCESSION	AC099763.1	GI:17017635	
VERSION			
KEYWORDS	HTG.		
SOURCE	Caenorhabditis briggsae		
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;		
REFERENCE	Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.		
AUTHORS	1 (bases 1 to 81624)		
TITLE	Washington University Genome Sequencing Center.		
JOURNAL	The C. briggsae Genome Sequencing Project		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 81624)		
TITLE	Waterston, R.		
JOURNAL	Direct Submission		
REFERENCE	Unpublished		
AUTHORS	3 (bases 1 to 81624)		
TITLE	Waterston, R.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (20-NOV-2001) Department of Genetics, Washington		
TITLE	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.		
JOURNAL	Louis, MO 63110, USA		
COMMENT	Submitted by:		
	Genome Sequencing Center		
	Department of Genetics, Washington University		
	St. Louis, MO 63110, USA		
	email: rtw@nemacode.wustl.edu		
NOTICE:	This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.		
	This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.		
FEATURES	Location/Qualifiers		
source	1..81624		
	/organism="Caenorhabditis briggsae"		
	/mol_type="genomic DNA"		
	/strain="Gujarat G16"		
	/db_xref="taxon:6238"		
ORIGIN			
Query Match	2.9%;	Score 39.6;	DB 3; Length 81624;
Best Local Similarity	58.5%;	Pred. No. 3.6;	
Matches	69;	Conservative 0;	Mismatches 49; Indels 0; Gaps 0;
QY	34	TCCAAACTGTGTAGTGCATGCAATGAGAAAGAAAGCGTGTGTGTGGCAGCTGAA	93
Db	8467	TTCAAAAATCTTATTATTATTCGACGCTGTTAAAAAAGAAAGTTTGTACCAATTGAG	8408
QY	94	GACGTGAAGAGAGACCAAGATAATTACACATGCGATACGTTGCATCAATGCTGTT	151
Db	8407	ACATGAATAATCAGTAAAAAATAATATGAGCTAAGACAAAGTACAAATGACTGTAGTT	8350

RESULT 13
 AC129421/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-62C13, *** SEQUENCING IN PROGRESS
 AC129421 290066 bp DNA linear HTG 20-NOV-2002
 *** 5 unordered pieces.
 AC129421
 AC129421.3 GI:25138103
 VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 290066)
 Muzny D, Marie, Metzger M, Lee, Abramson S, Adams C, Alder J, Allen C, Allen H, Alsbrooks S, Amin A, Angiano D, Anyaledechun V, Aoyagi A, Ayodeji M, Baca B, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Biewald K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant N, Buhay C, Burch P, Burrell K, Calderon E, Cardenas V, Carter K, Cavazos I, Caesar H, Center A, Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L, Davila M, Davis C, Davy-Carroll L, De Anda C, Dedertich D, Delgado O, Denison S, Deramo C, Ding Y, Dinh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Frazer C, Gabisl A, Ganta R, Garcia A, Garner T, Garza M, Gebregregis E, Geer K, Gill R, Grady M, Guerra W, Guevara M, Gunaratne P, Haaland M, Hamil C, Hamilton C, Hamilton K, Harrey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, Hodgson A, Hogues M, Hollins B, Howells S, Hulyk S, Hume J, Idlebird D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A, Karpathy S, Kelly S, Kelly S, Khan Z, King L, Kovar C, Kowalski C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J, Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J, Lounsbury L, Loulsegged H, Lozano R, Lu X, Ma Z, Maheshwari M, Mahindaratne M, Mahmoud M, Malloy K, Mangum A, Mangum B, Mapua P, Martin K, Martin R, Martinez E, Mawhinney S, McLeod M, McNeill T, Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Muidasa M, Murphy M, Nair L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nwakoeleneh O, Okunolu G, Olarnungson A, Pal S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Plankoch C, Plapper F, Polindexter A, Popovic D, Primus E, Pu L, Puozzo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs F, Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders W, Saverly G, Scherer S, Scott G, Shatsman S, Shen H, Shetty J, Shvartsbeyn A, Sison I, Sitter C, Sma's D, Sneed A, Sodergren E, Song X, Sorelle R, Soosa J, Steinle M, Strong R, Sutton A, Svatek A, Tabors P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Umanli K, Valas R, Vera V, Villasana D, Waldron L, Walker B, Wang J, Wang Q, Wang S, Warren J, Warren R, Wei X, White F, Williams G, Wilson R, Wleczky R, Woodan H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu F, Zhang J, Zhou J, Zhou X, Zhao S, Dunn D, von Niederhausen A, Weis R, Smith D, Holt R, Smith H, O, Weinstock G, and Gibbs R.A.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 290066)
 AUTHORS Worley K.C.
 JOURNAL Direct Submission
 Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 290066)
 AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission
 JOURNAL Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 20, 2002 this sequence version replaced gi:23267370. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GURV
 Center clone name: CH230-62C13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 216696 bases at least Q40
 Consensus quality: 216655 bases at least Q30
 Consensus quality: 221692 bases at least Q20
 Estimated insert size: 216915; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 40609: contig of 40609 bp in length
 * 40610 40709: gap of unknown length
 * 40710 281950: contig of 241241 bp in length
 * 281951 282050: gap of unknown length
 * 282051 283370: contig of 1320 bp in length
 * 283371 283470: gap of unknown length
 * 283471 286076: contig of 2606 bp in length
 * 286077 286176: gap of unknown length
 * 286177 290066: contig of 3890 bp in length.

FEATURES
 source 1..290066 Location/Qualifiers
 1..290066 "Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-62C13"
 32547..33462
 /note="clone_boundary"
 clone_end:5p6
 site:
 end_sequence:BH287162"
 40710..42036
 /note="wgs_contig"
 280810..281950
 /note="wgs_contig"

ORIGIN
 Query Match 2.9%; Score 39.6; DB 2; Length 290066;
 Best Local Similarity 52.4%; Pred. No. 3 7;
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

ORIGIN

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/60 mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES**Source**

```
1..186858
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosome="7"
  /map="7"
  /clone="RP24-34N22"
  /clone_lib="RPI-24"
  1334..1484
    /rpt_family="Alu"
  4834..4946
    /rpt_family="Alu"
  5035..5180
    /rpt_family="B2"
  6535..6651
    /rpt_family="L1"
  7259..7406
    /rpt_family="B2"
  7432..7567
    /rpt_family="Alu"
  7598..7663
    /rpt_family="B2"
  7664..7780
    /rpt_family="B4"
  7845..8055
    /rpt_family="B2"
  8235..8338
    /rpt_family="MER2_type"
  9419..9580
    /rpt_family="B4"
  10335..10537
    /rpt_family="B2"
  11210..11361
    /rpt_family="B4"
  11411..11497
    /rpt_family="Alu"
  11810..11962
    /rpt_family="B4"
  12045..12433
    /rpt_family="ERVK"
  12680..12727
    /rpt_family="ERV1"
  12774..12972
    /rpt_family="B2"
  13385..13502
    /rpt_family="Alu"
  13657..13857
    /rpt_family="B2"
  14075..14147
    /rpt_family="ID"
  14201..14226
    /rpt_family="B4"
  14227..14357
    /rpt_family="Alu"
```

```
repeat_region 14479..14860
/rpt_family="L1"
repeat_region 17823..17954
/rpt_family="Alu"
repeat_region 18495..18597
/rpt_family="Alu"
repeat_region 18545..18631
/rpt_family="BCL_MM"
repeat_region 18973..19078
/rpt_family="B4"
repeat_region 19119..19242
/rpt_family="Alu"
repeat_region 19243..19352
/rpt_family="B4"
repeat_region 19432..19499
/rpt_family="ID"
repeat_region 22725..22836
/rpt_family="B4"
repeat_region 23994..24053
/rpt_family="B4"
repeat_region 24034..24194
/rpt_family="B2"
repeat_region 24230..24487
/rpt_family="L1"
repeat_region 26457..26650
/rpt_family="B2"
repeat_region 28112..28276
/rpt_family="MER1_type"
repeat_region 29623..29699
/rpt_family="L2"
repeat_region 29973..30306
/rpt_family="ERVK"
repeat_region 30609..30705
/rpt_family="Alu"
repeat_region 31040..31148
/rpt_family="B4"
repeat_region 31927..31987
/rpt_family="Alu"
repeat_region 32001..32082
/rpt_family="L1"
repeat_region 32158..32257
/rpt_family="L1"
repeat_region 32837..32963
/rpt_family="Alu"
repeat_region 33965..34087
/rpt_family="B2"
repeat_region 34118..34211
/rpt_family="Alu"
repeat_region 34226..34383
/rpt_family="B4"
repeat_region 34746..34857
/rpt_family="Alu"
repeat_region 35012..35208
/rpt_family="B2"
repeat_region 36296..36543
/rpt_family="MIR"
repeat_region 37116..37236
/rpt_family="B4"
repeat_region 37323..37461
/rpt_family="B4"
repeat_region 37722..37853
/rpt_family="B4"
repeat_region 38215..38354
/rpt_family="Alu"
repeat_region 39131..39200
/rpt_family="Alu"
repeat_region 39392..39414
/rpt_family="B2"
repeat_region 39415..39496
/rpt_family="B4"
```

Query Match 2.9%; Score 39.4; DB 10; Length 186858;
Best Local Similarity 54.5%; Pred. No. 4.2;

	Matches	79;	Conservative	0;	Mismatches	66;	Indels	0;	Gaps	0;
OY	1215	ACGTAGATACATAGAGTCTATAGAAACATTAAGAGAGCTTGAAACCATTCATTCCTTA								1274
Db	130261	ATGTACAAACATATATCTCACACACAAATTAAGTAAATTGAAAAAGATTAATATCT								130202
OY	1275	GGGTCTCTCTTCTTCTGATCATCAAGAAATCATACACTCAACAGAACTTTCT								1334
Db	130201	TAAATCAAAACACACAGGTACACACACAAACATTAACAAAAACAAAAACA								130142
OY	1335	ATCTTCCTTATAGCAATTCCAAA								1359
Db	130141	ACCCTCCCTAAACAAAAA								130117

Search completed: February 9, 2005, 12:28:20
 Job time : 5980 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 08:54:14 ; Search time 752 Seconds
(without alignment)
10847.611 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtcgacgtgggtggtgatac.....accatcaatcaactaaca 1378

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001s:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1378	100.0	1378	5	AAf79834
2	48.2	3.5	2000	8	ADA71938
3	40	2.9	6197	6	ABN80257
4	37.2	2.7	2000	8	ADA71938
5	36.2	2.6	2672	6	ADA29004
6	36.2	2.6	2672	12	ADN96821
7	36.2	2.6	4590	5	AAH24065
8	35.8	2.6	134738	11	ACN44182
9	35.4	2.6	993	4	AAf81969
10	35.4	2.6	993	10	ADBE2962
11	35	2.5	6971	6	ABL33237
12	35	2.5	207433	5	ABE72040
13	35	2.5	207433	8	ABX74891
14	35	2.5	207433	12	ADJ36614
15	35	2.5	207433	12	ADL81193
16	34.8	2.5	4416	13	ADS61188
17	34.6	2.5	5059	2	AAH84332
18	34.4	2.5	2799	10	ABV74497
19	34.4	2.5	8805	6	ABK40016
20	34.4	2.5	32874	9	ADA02648

21	34.4	2.5	32874	10	ADB72386
22	34.4	2.5	32874	10	ADE95896
23	34.2	2.5	1500	10	ADC92243
24	34.2	2.5	90435	12	ADO59524
25	34	2.5	6283	6	ABK39991
26	34	2.5	6283	6	ABK39991
27	34	2.5	339234	12	ADQ59437
28	33.8	2.5	648	10	ADQ56087
29	33.8	2.5	980	10	ADK59677
30	33.8	2.5	103471	12	ADO97668
31	33.6	2.4	25179	13	ABD33248
32	33.6	2.4	191284	12	ADQ97957
33	33.4	2.4	403	4	AAI36496
34	33.4	2.4	403	4	ABA26537
35	33.4	2.4	1781	2	AAV35130
36	33.4	2.4	4206	2	AAV35115
37	33.4	2.4	4792	2	AAH83004
38	33.4	2.4	5058	2	AAV35114
39	33.4	2.4	6476	2	AAH24302
40	33.4	2.4	7189	4	AAH30652
41	33.4	2.4	7189	4	AAH28725
42	33.4	2.4	7189	8	ACA03415
43	33.4	2.4	7189	9	ADB96763
44	33.4	2.4	7189	10	ADG41921
45	33.4	2.4	7189	11	ADI97695

ALIGNMENTS

RESULT 1	AAf79834	standard; DNA; 1378 BP.
ID	AAf79834	
XX	AAf79834;	
AC	AAf79834;	
XX		
DT	30-MAY-2001 (first entry)	
XX		
DE	Mycelia sterilia promoter sequence.	
XX		
KW	Promoter; terminator; regulatory region; filamentous fungus;	
KW	protein production; ds.	
OS	Mycelia sterilia.	
XX		
PN	WO200118219-A1.	
XX		
PD	15-MAR-2001.	
XX		
PF	07-SEP-2000; 2000WO-JP006104.	
XX		
PR	07-SEP-1999; 990P-00252851.	
XX		
PA	(MEIJU) MEIJU SEIKA KAISHA LTD.	
XX		
PI	Watanabe M., Murakami T;	
XX		
DR	WPI; 2001-235202/24.	
XX		
PT	New promoter and terminator functioning synchronously, useful for	
PT	regulating expression of endogenous gene in filamentous fungi,	
PT	particularly transformed Mycelia sterilia capable of producing high	
PT	yields of target protein/substance.	
XX		
PS	Claim 2; Page 18-20; 26pp; Japanese.	
XX		
CC	The present invention provides the sequences of a promoter and terminator	
CC	from Mycelia sterilia which can be used to regulate the expression of an	
CC	endogenous gene in filamentous fungi, particularly transformed Mycelia	
CC	sterilia, to produce useful target proteins or other substances. The	
CC	present sequence is the promoter of the invention	
XX		
SQ	Sequence 1378 BP; 396 A; 289 C; 340 G; 353 T; 0 U; 0 Other;	

Query Match	100.0%	Score 1378;	DB 5;	Length 1378;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 1378;	Conservative 0;
Matches 1378;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1	GTCCGCGGGGGTGGATATCATGCTGGTGTCCAAAACGTTGTTAGTCATGAT	60	
Db	1	GTCCGCGGGGGTGGATATCATGCTGGTGTCCAAAACGTTGTTAGTCATGAT	60	
OY	61	GAGGAAGAAACGCTGTGTGTGTGCGAGCTGAAGA CTGAAGAGAGCCAAAGATATTC	120	
Db	61	GAGGAAGAAACGCTGTGTGTGTGCGAGCTGAAGA CTGAAGAGAGCCAAAGATATTC	120	
OY	121	ACAATGCGATACGCTTGCATCATATGCTTGTTCAGAGACGTTGCATCTACCTGGTGT	180	
Db	121	ACAATGCGATACGCTTGCATCATATGCTTGTTCAGAGACGTTGCATCTACCTGGTGT	180	
OY	181	TCCCGCTTCTGTTGACAAAGTACAGTATCCGATGACACCCCGCGAAGGAAATCCG	240	
Db	181	TCCCGCTTCTGTTGACAAAGTACAGTATCCGATGACACCCCGCGAAGGAAATCCG	240	
OY	241	GAGTTCAAAAGAGGTGTGCTCTACGCGCA TTTAGTATAGATGAGCATAGGGTTGACGTAA	300	
Db	241	GAGTTCAAAAGAGGTGTGCTCTACGCGCA TTTAGTATAGATGAGCATAGGGTTGACGTAA	300	
OY	301	GCTGAAGCTGATTACGAGACATGAGACAGAAATACACGCTTGTATGCGTTCCGCT	360	
Db	301	GCTGAAGCTGATTACGAGACATGAGACAGAAATACACGCTTGTATGCGTTCCGCT	360	
OY	361	GCTTACTTAAGTATATCCAAAGACACACAGCCGAAAGAAACCGATGCTGTAGAGGG	420	
Db	361	GCTTACTTAAGTATATCCAAAGACACACAGCCGAAAGAAACCGATGCTGTAGAGGG	420	
OY	421	TTCCCTTTAGAGCTACATGCTTACGCTGATGATAGAAACATCAATGCGCAATCAAGTT	480	
Db	421	TTCCCTTTAGAGCTACATGCTTACGCTGATGATAGAAACATCAATGCGCAATCAAGTT	480	
OY	481	AGTATACCTGACGCTACATCGCTTTCTCCGATCTTGCCTTAAATATATGCTGCTGTC	540	
Db	481	AGTATACCTGACGCTACATCGCTTTCTCCGATCTTGCCTTAAATATATGCTGCTGTC	540	
OY	541	GAACTGCGGTACTGCTTACTACTACTGTTCTCCGTTGAAGCTCTAGAGACAGCGCG	600	
Db	541	GAACTGCGGTACTGCTTACTACTACTGTTCTCCGTTGAAGCTCTAGAGACAGCGCG	600	
OY	601	CGTTTGTAGACCTACATGATGCGACATCTTAAGAGAGGATCTGAGACATTTCTTAAGGC	660	
Db	601	CGTTTGTAGACCTACATGATGCGACATCTTAAGAGAGGATCTGAGACATTTCTTAAGGC	660	
OY	661	ATCCATATAGGCAATGCGGCTTAAGTCCGCAATGGAAGAGATPAAGGGGGGTGTGAAGTG	720	
Db	661	ATCCATATAGGCAATGCGGCTTAAGTCCGCAATGGAAGAGATPAAGGGGGGTGTGAAGTG	720	
OY	721	GTGCTGCAAAAAGAGAGTGAATGCGCTATPCCAGCCGCTPAAGAGAGGGGCTGACAGCTGT	780	
Db	721	GTGCTGCAAAAAGAGAGTGAATGCGCTATPCCAGCCGCTPAAGAGAGGGGCTGACAGCTGT	780	
OY	781	CTGACAGCTGTGAATACGTCATCTTGTGATGTATGTCACCTAATGTGACAGATGCAAA	840	
Db	781	CTGACAGCTGTGAATACGTCATCTTGTGATGTATGTCACCTAATGTGACAGATGCAAA	840	
OY	841	TGCTGATTTGGTTAAATATGCGCATGTATGATGAGTGCAGAAAACAGTTTAAATCTAAGTT	900	
Db	841	TGCTGATTTGGTTAAATATGCGCATGTATGATGAGTGCAGAAAACAGTTTAAATCTAAGTT	900	
OY	901	AAAGGAAGCTGAAGCTGAACCTGTCAGAAATTAAGCCCTGTGGAATTAAGAGTTGATTA	960	
Db	901	AAAGGAAGCTGAAGCTGAACCTGTCAGAAATTAAGCCCTGTGGAATTAAGAGTTGATTA	960	
OY	961	CCCAATTACGTGTCAAGAGGTGTCTCATATGTCTGAGAGTTCCCTGTGCGCATTTGTGGGT	1020	
Db	961	CCCAATTACGTGTCAAGAGGTGTCTCATATGTCTGAGAGTTCCCTGTGCGCATTTGTGGGT	1020	

OY	1021	AACTATTTTCATAGTGGGCGAGAAATGCACCTCTATTTTCAATTTGAATCTAAATATTCGG	1080	
Db	1021	AACTATTTTCATAGTGGGCGAGAAATGCACCTCTATTTTCAATTTGAATCTAAATATTCGG	1080	
OY	1081	GTAGAGATTTCTCAATAGTGTCTTCTGCTGTCACTTACACATCAATGCGGGTCAACAAGT	1140	
Db	1081	GTAGAGATTTCTCAATAGTGTCTTCTGCTGTCACTTACACATCAATGCGGGTCAACAAGT	1140	
OY	1141	ATACAGCTTCAATPAGAGAGTGGGCGATTTGAGTACGATCCGATCGAACCCGGAACGGTT	1200	
Db	1141	ATACAGCTTCAATPAGAGAGTGGGCGATTTGAGTACGATCCGATCGAACCCGGAACGGTT	1200	
OY	1201	CAAGCATGGGCGTACGATGATCATATAGATCATATGAAACATPAAGAGAGCTTGAAGAAC	1260	
Db	1201	CAAGCATGGGCGTACGATGATCATATAGATCATATGAAACATPAAGAGAGCTTGAAGAAC	1260	
OY	1261	CATTCAATTCCTTAAGGGGTCTCTTCTTCTTGTGATCATCATCAAGATCATCACTCAAC	1320	
Db	1261	CATTCAATTCCTTAAGGGGTCTCTTCTTCTTGTGATCATCATCAAGATCATCACTCAAC	1320	
OY	1321	CAGGAACCTTTTCTATCTTCCCTATPAGCAATTCGCAAAACCCATCAATCAACTTACA	1378	
Db	1321	CAGGAACCTTTTCTATCTTCCCTATPAGCAATTCGCAAAACCCATCAATCAACTTACA	1378	

RESULT 2
ADA71938
ADA71938 standard; DNA; 2000 BP.

XX AC ADA71938;
XX DT 20-NOV-2003 (first entry)
XX DE Rice gene, SEQ ID 5263.
XX KW Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.
XX OS Oryza sativa.
XX PN WO200300898-A1.
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
XX DR MPI; 2003-175290/17.
XX PT Identifying at least one gene involved in plant resistance or response to
XX PT pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX PT gene expression.
XX PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX CC The present invention relates to a method (M1) for identifying genes
XX CC involved in plant resistance or response to pathogenic infection. M1
XX CC comprises identifying a gene whose expression is significantly altered in
XX CC the incompatible interaction of plant gene expression relative to
XX CC expression of the gene in an uninfected plant, in a mutant plant that
XX CC does not express a gene associated with response to pathogenic infection,
XX CC or in a corresponding incompatible or compatible interaction. (M1) is
XX CC useful for conferring resistance to resistance or tolerance to a plant to
XX CC bacterial, fungal or viral infection. The present sequence was used to
XX CC illustrate the invention.

Seq	Sequence	2000 BP	336 A	265 C	284 G	363 T	0 U	752 Other
Query Match	3.5%;	Score 48.2;	DB 8;	Length 2000;				
Best Local Similarity	10.0%;	Pred. No. 0.0034;						
Matches	79;	Conservative 339;	Mismatches 372;	Indels 1;	Gap 1;			
Qy	585	CCTAGGACAAGGCGCGCTTTGTGTATACCTCACTGATGTCACACTCTTTAAAGCAGGATCTCG	644					
Db	28	SCARGSSSRMSKMSKMSKRRKSSCGKCKMTTRKSKMSYSSASGRTGSWSSGYSYG	87					
Qy	645	AGACATTTTCTAAGGCATCCATATATGAGCTTGGGGGCTAATGCGCATTTGAAGAGATTA	704					
Db	88	KGMKRRYKRSKRRRGRGRGRMRSRMRMRGRTRRCARSRMAAGSGRRMGSRMSYMM	147					
Qy	705	GAGGGGTGAAAGTGTGTGTCAAAAGGAGGTCATATGGCTATATACAGCCGCTAAGCAG	764					
Db	148	CYARGCGSKRRKSKSGWGTCTRGARGGSGWSGAKYKSSGMSRBMMSCGRSGGR	207					
Qy	765	GTGGGCTTGACAGCTGTCTGCAGCTGTGAATACGTCACTGTTAGTATGTCCACTTA	824					
Db	208	RSAYSRYYGTSTKRYGTYYKMTYTSARCFRAWTTTSYVACSSTYWCBSKRSRSMWKMR	267					
Qy	825	TGTACGACGATGCAAAATGCTGATTTGGGTTAAATAGGACATGATGATGTGTCGCCAAAC	884					
Db	268	KMRSRSTSGWTSWSYROMMYCTATKYSYSRWCTMYTRGGGWRATATYWGIVSRMAWTK	327					
Qy	885	ACGTTTATGATCTAGTTTAAAGGAACTGAAAGCTGAACCTGTCAGAAATTAACCTGTTGG	944					
Db	328	KMYWRYGKMGMRGMMAGMRMRMSRCRMSKACYWRMRMRMTTRRRRAKKSSRTSRKX	387					
Qy	945	AATACACGTTGATTAACCAATTCAGTCTCAAGGGTGTCTGATATATGCTGAGCTTCCC	1000					
Db	388	RKMCRKRKKYKMRGSRMRSCBRAMMRCSRSGAMWKGCRCMCRCMKSYGMRRKMSW	447					
Qy	1005	TGTGCATTTGGGGTAACTATTTGATAGTGGGGCAGATGCACTATTTTCAATTGA	1066					
Db	448	KMAASKYKMSHMYWRKKKCSRTTMMGKTGAGMGMTGRCHRYKRSRGMKRCRRRRMR	507					
Qy	1065	ATCTAACTATCTGGGTAGAGTTCTCAATGCTCTTCGCGTCACTTACACATCA	1124					
Db	508	MYRRMRKRYMSARITMYCAKAKYSYSAARARCMYRKGYVMAGMMMKRYKMYMYM	567					
Qy	1125	TGGGGTCAACAACGTATACAGCTTCA-TAGAGAGTGGCGCATTAAGTAGTACCGCAT	1188					
Db	568	MMWYRKYSKSGWYCMYSYASCMKSAARAAGKMRKSRMSMSRSCRKRCASRK	627					
Qy	1184	CGAACCCGGAAGCGGTTCAAGACATAGGGGCTACGTAGATACATAGAGTATGAAACATA	1244					
Db	628	SSAKRYAMWGMTTSSRMSRMSLYTCYMKWSMSSTCYMMYSRYTAYAKGSYWRYY	687					
Qy	1244	AAAGAGGCTTGAAGAACCATTCAAAATCCTAAGGTCCTCTCTTCTCTCATCATCA	1300					
Db	688	RAWCMYMRMYRRYRSTYWTYMAWTTSTRAMATGKYSGRYWTSMYKCKCKSMKTRSM	747					
Qy	1304	GAATCATACACTCAACACAGAGACTCTTCTATCTTCCTATAGACATTTCCAAAAACCA	1360					
Db	748	YYWSWMAKTMGMWRRYATRRMMMYRYSMKYTYCTCMGYMMWMTYTRRRYMYKYCT	807					
Qy	1364	TCAATCAACCT 1374						
Db	808	KTYWTSATYTW 818						
RESULT 3								
ABN80257/c								
ID	ABN80257	standard; DNA; 6197 BP.						
XX	ABN80257;							
XX	15-JUL-2002	(first entry)						
XX	Human	chemically modified disease associated gene	SEQ ID NO 274.					
XX								

[illegible]

[illegible]

QY		722	TGTGTCAAAAGAGAGTGCATTGGCTATACCGCCCTTAAGCAGTGCGGCTAC	773
Dd		199	SSKMKTKMSKSGSMRWTCSSWCSCCYTCYYGAMCWSGCCMMYMGSGCYTRG	148
			RESULT 5	
		AAD29004/C		
ID		AAD29004	standard; DNA; 2672 BP.	
XX				
XX		AAD29004;		
XX				
DT		07-MAY-2002	(first entry)	
XX				
DE		Bugula neritina cosmid clone 5B Pst D4/Cl overlap DNA sequence.		
XX				
KW		polyketide; bryopyran ring; byrostatin; breast cancer; anticancer;		
XX		antifungal; antimicrobial; immunomodulatory; polyketide synthase; PKS;		
XX		enzyme; ds.		
OS				
XX		Bugula neritina.		
PN		WO200111024-A2.		
XX				
PD		15-FEB-2001.		
XX				
Pf		04-AUG-2000; 2000MO-US021326.		
PR				
PA		04-AUG-1999; 99US-0147283P.		
XX		(REGC) UNIV CALIFORNIA.		
PI		Haygood M, Davidson SK, Allen SW, Hildebrand M;		
DR		WPI; 2002-154285/20.		
PT				
PT		Composition comprising a polypeptide isolated from marine organism, which		
PT		catalyzes at least one step in synthesis of polyketide/bryopyran ring,		
PT		useful for producing polyketide or bryopyran ring containing		
compositions.				
PS				
XX		Claim 64; Fig 17B; 233pp; English.		
XX				
CC		The invention relates to compositions comprising nucleic acid molecules		
CC		encoding a polypeptide which catalyzes at least one step in synthesis of		
CC		polyketides including bryopyran ring, such as byrostetins. These novel		
CC		sequences are derived from marine organisms. Compositions containing		
CC		sequences of the invention are useful for producing base structure,		
CC		bryopyran rings that can form the basis of combinatorial chemistry to		
CC		form a wide variety of compounds which can be screened for bioactivities		
CC		including anticancer activity. The cloned genes and linked genes involved		
CC		in byrostatin synthesis can be used to screen environmental samples for		
CC		polyketide synthase (PKS) genes. They are also used for combinatorial		
CC		creation of novel polyketide/byrostatin analogues that may exhibit		
CC		improved anti-cancer properties. Compositions of the invention are useful		
CC		for producing byrostatin and its analogues which are useful for treating		
CC		breast cancer and as anticancer, antifungal antimicrobial and		
CC		immunomodulatory compounds. They are useful for producing novel		
CC		polyketides such as bryopyran rings including byrostetins. The present		
CC		sequence is Bugula neritina cosmid clone 5B Pst D4/Cl overlap DNA		
CC		sequence used in the exemplification of the invention		
XX				
SQ		Sequence 2672 BP; 666 A; 502 C; 570 G; 660 T; 0 U; 254 Other;		
Query Match		2.6%; Score 36.2; DB 6; Length 2672;		
Best Local Similarity		24.1%; Pred. No. 3.4;		
Matches		41; Conservative 60; Mismatches 69; Indels 0; Gaps 0		
QY		674 TTGGGCGGCTAAGTCGCAATGAAGAAGATAGGGGGGTGAAAGGCTGTCAAAGAAG	733	
Dd		702 TTCWGCGCRRTCKASMTKYTKKTWTCCAAAGSGSRGSYKTKMTWTYSGMANCRG	643	
QY		734 AGGTGATTTGGTATACCAGCCGCTAACGAGGTGGCTAGCAGCTGTCTCAGACTGTGAA	793	

Db 642 MYSMGTGTTTCRAAMWKSNGYTKCGTGTAGTGYWCCAMWWSCRMWMSCRGSGSKSCTTKMT 583
Oy 794 TAACGTCACTTGCTTAGTATGTCCACTAATGTACAGACGATGCAAAATGC 843
Db 582 KAAMRAMMCTYYMTTSGAAMAASYTKYKAGGKMGKGGKCMCRAMTKY 533

RESULT 6

ADN96821/c
ID ADN96821 standard; DNA, 2672 BP.

AC ADN96821;

DT 01-JUL-2004 (first entry)

XX Bugula neritina PKS cosmid clone 5B Pst D4/C1 overlap DNA SegID 34.

XX ds; degenerate; polyketide; bryopyran ring; bryostatin;

KW polyketide synthase; PKS; toxic; cytostatic; immunomodulatory;

KM protein therapy; cancer; metastasis.

XX Bugula neritina.

PN W02003099219-A2.

PD 04-DEC-2003.

XX 20-MAY-2003; 2003WO-US016299.

XX 20-MAY-2002; 2002US-0382181P.

PA (REGC) UNIV CALIFORNIA.

PA (HAYG/) HAYGOOD M.

PA (HILD/) HILDEBRAND M.

PA (ANDE/) ANDERSON C.

PA (WAGG/) WAGGONER L E.

PA (SHER/) SHERMAN D H.

PA (LIUH/) LIU H.

XX Haygood M, Hildebrand M, Anderson C, Waggoner LE, Sherman DH;

PI Liu H;

XX WPI; 2004-053143/05.

XX New compositions comprising a polyketide synthase or nucleic acid

FT encoding the polyketide synthase, useful in biosynthesizing polyketides,

PT bryopyran rings and bryostatins having anti-cancer or antimetastatic

PT activity.

XX Claim 13; SEQ ID NO 34; 342pp; English.

XX This invention relates to a novel composition that comprises at least one

CC polypeptide that catalyzes the one step synthesis of a polyketide or

CC bryopyran ring. Specifically, it refers to nucleic acid molecules derived

CC from marine organisms that encode enzymes that catalyze the synthesis of

CC bioactive compounds such as polyketides and bryostatins that are based on

CC the use of polyketide synthases (PKSs) to generate toxic polyketides that

CC exhibit cytostatic and immunomodulatory activities, such that they can be

CC used for protein therapy in the treatment of cancer and metastasis. This

CC polynucleotide is a Bugula neritina PKS DNA sequence of the invention.

XX Sequence 2672 BP; 686 A; 502 C; 570 G; 660 T; 0 U; 254 Other;

SO Query Match 2.6%; Score 36.2; DB 12; Length 2672;

Best Local Similarity 24.1%; Pred. No. 3.4;

Matches 41; Conservative 60; Mismatches 69; Indels 0; Gaps 0;

Oy 674 TTGGGCGCTAGTCGCGCATTTAGAGATAGAGGGGGGTGTAAGTGTGTCTCAAAAGC 733

Db 702 TTCWGMGCMRSTCAKSKWYKTKWTWCCAGAGSTGRKSKYKGTWTKWTYSGWMANCCRG 643

Oy 734 AGTCGATTTGCTATACAGCCGCTAAGCAGGTGGCTAGACGCTGTCTGCAGCTGTGAA 793
Db 642 MYSMGTGTTTCRAAMWKSNGYTKCGTGTAGTGYWCCAMWWSCRMWMSCRGSGSKSCTTKMT 583
Oy 794 TAACGTCACTTGCTTAGTATGTCCACTAATGTACAGACGATGCAAAATGC 843
Db 582 KAAMRAMMCTYYMTTSGAAMAASYTKYKAGGKMGKGGKCMCRAMTKY 533

RESULT 7

AAH24065
ID AAH24065 standard; DNA, 4590 BP.

AC AAH24065;

DT 29-AUG-2001 (first entry)

XX Yeast AOD9604-associated DNA sequence, SEQ ID NO:1.

XX Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;

KW modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;

KM lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;

KW functional food; transgenic yeast; fat/lean ratio; food use; ds.

XX Saccharomyces cerevisiae.

XX Key location/Qualifiers

FT misc_feature 10

FT /tag= a

FT /note= "Represented as * in the specification"

FT misc_feature 3617

FT /tag= b

FT /note= "Represented as * in the specification"

FT misc_feature 3649

FT /tag= c

FT /note= "Represented as * in the specification"

FT misc_feature 3679

FT /tag= d

FT /note= "Represented as * in the specification"

FT misc_feature 3819

FT /tag= e

FT /note= "Represented as * in the specification"

FT misc_feature 3862

FT /tag= f

FT /note= "Represented as * in the specification"

FT misc_feature 3864

FT /tag= g

FT /note= "Represented as * in the specification"

FT misc_feature 3888

FT /tag= h

FT /note= "Represented as * in the specification"

FT misc_feature 3890

FT /tag= i

FT /note= "Represented as * in the specification"

FT misc_feature 3912

FT /tag= j

FT /note= "Represented as * in the specification"

FT misc_feature 3914

FT /tag= k

FT /note= "Represented as * in the specification"

FT misc_feature 3938

FT /tag= l

FT /note= "Represented as * in the specification"

FT misc_feature 3939

FT /tag= m

FT /note= "Represented as * in the specification"

FT misc_feature 3941

FT /tag= n

FT /note= "Represented as * in the specification"

FT misc_feature 3943

FT /tag= o

FT /note= "Represented as * in the specification"

FT misc_feature 4361

FT /tag= p

FT /note= "Represented as * in the specification"

QY 482 GTATACCTGACGCTACATC 500
Db 118373 GTACAGCTGACCCCTTCAC 118355

RESULT 9
AAF83969/c
ID AAF83969 standard; DNA; 993 BP.
AC AAF83969;
XX
XX 06-AUG-2001 (first entry)
DT
XX
XX Human SER4 nucleotide sequence.
DE
XX Serine/threonine kinase-like protein; SERX; breast; testicular; cancer;
KM cardiovascular; autoimmune disorder; cytostatic; cardiac;
KM gene therapy. immunosuppressive; antiaesthetic; antifibrinolytic; SER4;
KM ds.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..993
FT /*tag= a
FT /product= "SER4"
XX
XX MO200136645-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000MO-US031744.
XX
XX 17-NOV-1999; 99US-0165986P.
PR 09-FEB-2000; 2000US-0181347P.
PR 03-APR-2000; 2000US-0194195P.
PR 05-APR-2000; 2000US-0194839P.
PR 07-APR-2000; 2000US-0195637P.
PR 13-APR-2000; 2000US-0197080P.
PR 15-SEP-2000; 2000US-0232677P.
PR 16-NOV-2000; 2000US-00715427.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Quinn KE, Spytek KA, Majumder K, Vernet C, Burgess C;
PI Fernandes E, Taupier R, Rastelli L, Herrmann JL;
XX
XX WPI; 2001-336006/35.
DR P-PSDB; AAB85038.
XX
XX New isolated SERX polynucleotides and polypeptides related to the
PT serine/threonine kinase family of proteins, useful for determining and
PT treating diseases related to altered levels of the polypeptide.
XX
XX Claim 9; Page 21; 127pp; English.

CC The invention relates to serine/threonine kinase-like proteins, termed
CC SERX. The SERX polynucleotides can be used to express SERX protein (e.g.
CC via a recombinant expression vector in a host cell in gene therapy
CC applications), to detect SERX mRNA (e.g. in a biological sample) or a
CC genetic lesion in a SERX gene and modulate SERX activity. The SERX
CC polypeptides, polynucleotides, antibodies can also be used to treat or
CC prevent a pathology associated with SERX. Disorders associated with or
CC aberrant SERX expression or activity include breast or testicular cancer,
CC cardiovascular defect, autoimmune disorders and asthma. Disorders
CC associated with aberrant SERX-8 expression include, blood clotting
CC hematopoietic and tumor related disorders. The present sequence
CC represents the nucleotide sequence of human SER4
XX
XX Sequence 993 BP; 229 A; 305 C; 264 G; 195 T; 0 U; 0 Other;

Query Match 2.6%; Score 35.4; DB 4; Length 993;
Best Local Similarity 54.1%; Pred. No. 3.6;

Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 904 GGAAGCTGAAGCTGAACCTGTGAAATGAAGCTGTGGAAATCAAGTTGATTAACCC 963
Db 708 GGAAGCTCAAGGCTGTATCCGTCAGATGAGAGGCTGCTGAGCCGAGGCAATGATTTCTC 649
QY 964 AATTCAGTCGTCAAGGAGTGTCTGTATGCTGAGCTTCCCTGTGCAATGTGGGTTAC 1023
Db 648 AGTTGTCCCTCGAGAGGTGATACATCAACCGCAGCCTCGAAGGCCCTTGTAGGCTGGC 589
QY 1024 TATTTCAATGTCG 1036
Db 588 CAGCTCGTAGAGG 576

RESULT 10
ADE29362/c
ID ADE29362 standard; cDNA; 993 BP.
AC ADE29362;
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human cDNA encoding serine/threonine kinase SER4.
DE
XX
XX Human; ss; gene: serine threonine kinase; SER4;
KM blood coagulation disorder; liver disorder; hepatoma cell disorder;
KM hepatocellular cell disorder; obesity; diabetes; cancer; gene therapy.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..993
FT /*tag= a
FT /product= "SER4"
XX
XX US2003077697-A1.
XX
XX 24-APR-2003.
XX
XX 03-JUL-2001; 2001US-00898837.
XX
XX 17-NOV-1999; 99US-0165986P.
PR 09-FEB-2000; 2000US-0181347P.
PR 03-APR-2000; 2000US-0194195P.
PR 05-APR-2000; 2000US-0194839P.
PR 07-APR-2000; 2000US-0195637P.
PR 13-APR-2000; 2000US-0197080P.
PR 03-JUL-2000; 2000US-0215906P.
PR 15-SEP-2000; 2000US-0232677P.
XX
XX (GERL/) GERLACH V L.
PA (MACD/) MACDOUGALL J R.
PA (QUIN/) QUINN K E.
PA (MAJU/) MAJUMDER K.
PA (SPYT/) SPYTEK K A.
PA (VERN/) VERNET C.
PA (BURG/) BURGESS C E.
PA (FERN/) FERNANDES E R.
PA (RAST/) RASTELLI L.
PA (HERR/) HERRMANN J L.
PA (SPAD/) SPADERNA S K.
PA (SHIM/) SHIMKETS R A.
PA (TAUP/) TAUPIER R J.
XX
XX Gerlach VV, MacDougall JR, Quinn KE, Majumder K, Spytek KA;
PI Vernet C, Burgess CE, Fernandes ER, Rastelli L, Herrmann JL;
PI Spaderna SK, Shimkets RA, Taupier RJ;
XX
XX WPI; 2003-615999/58.
DR P-PSDB; ADE29363.
XX
XX New serine/threonine protein-kinase like polypeptides and genes, useful

PT in gene therapy, diagnosis or prognosis of e.g. diseases of blood
PT coagulation, obesity; diabetes, or cancers.
PS Claim 9; SEQ ID NO 8; 77pp; English.
XX
CC The invention relates to a new isolated polypeptide has a sequence
CC (designated SER1, SER2, SER4, SER5 and SER6, respectively) fully defined
CC in the specification (or their mature forms, variants or fragments). Also
CC included are an isolated nucleic acid molecule (comprising: a nucleic
CC acid sequence encoding any of the SER polypeptides above; a nucleic acid
CC fragment encoding at least a portion of the SER polypeptides or their
CC variants; or their complements), vectors comprising the nucleic acid
CC molecules, cells comprising the vectors, an antibody that binds
CC immunospecifically to the SER polypeptide, determining the presence or
CC amount of the SER polypeptide or nucleic acid in a sample, identifying an
CC agent that binds to the SER polypeptide, identifying a (potential)
CC therapeutic agent for use in the treatment of a pathology related to
CC aberrant expression or aberrant physiological interactions of the SER
CC polypeptide, modulating the activity of the SER polypeptide,
CC pharmaceutical compositions comprising a carrier (and the SER
CC polypeptide, nucleic acid molecule or antibody), kits comprising in one
CC or more containers any of the compositions above, screening for a
CC modulator of activity or latency or predisposition to a pathology
CC associated with the SER polypeptide and determining the presence of or
CC predisposition to a disease associated with altered levels of the SER
CC polypeptide or nucleic acid molecule. The SER polypeptide, nucleic acid
CC or anti-SER antibody is useful for treating or preventing a pathological
CC state or pathology associated with the SER polypeptide in a subject,
CC particularly a human. These pathologies include diseases of blood
CC coagulation, human liver, hepatoma cells or hepatocellular cells,
CC obesity, diabetes, or cancers. The polypeptide or polynucleotide is also
CC useful in gene therapy, diagnostics, prognosis, or research. The present
CC sequence encodes SER4.
XX
SQ Sequence 993 BP; 229 A; 305 C; 264 G; 195 T; 0 U; 0 Other;
Query Match 2.6%; Score 35.4; DB 10; Length 993;
Best Local Similarity 54.1%; Pred. No. 3.6;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 904 GGGAGCGTGAAGCTGTCAGAAATAGAGCTGTGGAATACAGTTGATTAACCC 963
DB 708 GCGAAGCTCAGAGGCTGATCCGTCAGATGAGAGGCTGCGCGAGGAGATTTTTC 649
QY 964 AATTCACTGTCAGAGGCTGCTGATATGCTGAGCTTCCTGTGCAATTGGGGTAAAC 1023
DB 648 AGTTGTCCCTCGAGAGGTCATACACTACCGCAGACCTCGAAAGCCCTTGAAGGCTGGC 589
QY 1024 TATTTTCATATGG 1036
DB 588 CAGCTCGTAGAG 576
RESULT 11
ABL3237/c
ID ABL3237 standard; DNA; 6971 BP.
XX
AC ABL3237;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1210.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytoskeletal; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineuritic; antidiabetic; antidiabetic; antiporiatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.
XX
XX Homo sapiens.
OS

XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 1210; 32pp + Sequence listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 6971 BP; 1665 A; 242 C; 1831 G; 3232 T; 0 U; 1 Other;
Query Match 2.5%; Score 35; DB 6; Length 6971;
Best Local Similarity 53.2%; Pred. No. 14;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 1238 AACATAAAGAGCTTGAAGAACCATTCATTAAGGCTCTCTTTCTTGCAATCA 1297
DB 2068 AAAAAAAAAAACTATATAAAACCTTAACCAAGACTCTCTTACTTCTTAACA 2009
QY 1298 CATCAAGATCATCTACTCAACACGAGACTCTTCTATCTTCCCTATAGCAATTTCCCA 1357
DB 2008 AACTATATATCTATTAACAAATCCCAAAAAAAAACTTAACTCCCTCACTCCACCCCA 1949
QY 1358 AACCATCAATCAACTAA 1376
DB 1948 CAACATCCCCCAACTAAA 1930
RESULT 12
AB272040/c
ID AB272040 standard; DNA; 207433 BP.
XX
AC AB272040;
XX
DT 03-APR-2003 (first entry)
XX
DE Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX
KW Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
KW antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
KW obesity; inflammatory bowel disease; promoter; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200178894-A2.
XX
XX 25-OCT-2001.
PD
XX 13-APR-2001; 2001WO-US012245.
XX
XX 13-APR-2000; 2000US-00548797.
PR

XX	(GENO-) GENOME THERAPEUTICS CORP.	
PA		
XX	Keith T;	
P1		
XX	WPI; 2001-639428/73.	
DR	P-PSDB; ABR00926.	
XX		
PT	Isolated genes (Gene 216) from human chromosome 20p13-p12 and the	
PT	proteins they encode, useful for the prevention, diagnosis and treatment	
PT	of asthma, obesity and inflammatory bowel disease.	
XX		
PS	Example 4; Fig 7; 520pp; English.	
XX		
CC	The invention relates to isolated genes (Gene 216) from human chromosome	
CC	20p13-p12 and the proteins they encode. The nucleic acids and proteins	
CC	may be used in the prevention, diagnosis and treatment of diseases	
CC	associated with inappropriate Gene 216 expression. For example, the	
CC	nucleic acids (or vectors) and proteins may be used to treat disorders	
CC	associated with decreased expression by rectifying mutations or deletions	
CC	in a patient's genome that affect the activity of gene 216 by expressing	
CC	inactive proteins or to supplement the patients own production of Gene	
CC	216 proteins. Additionally, the nucleic acids may be used to produce the	
CC	secreted Gene 216 protein, by inserting the nucleic acids into a host	
CC	cell and culturing the cell to express the protein. The nucleic acids and	
CC	complementary sequences may also be used as DNA probes in diagnostic	
CC	assays to detect and quantitate the presence of similar nucleic acid	
CC	sequences in samples and therefore which patients may be in need of	
CC	restorative therapy. The Gene 216 protein may also be used as antigens in	
CC	the production of antibodies against Gene 216 and in assays to identify	
CC	modulators of Gene 216 expression and activity. The anti-Gene 216	
CC	antibodies and antagonists may also be used to down regulate expression	
CC	and activity. The anti-Gene 216 antibodies may also be used as diagnostic	
CC	agents for detecting the presence of Gene 216 proteins in samples (e.g.	
CC	by enzyme linked immunosorbent assay or ELISA). Disorders that may be	
CC	prevented, diagnosed and/or treated by the above methods include, for	
CC	example asthma, obesity and inflammatory bowel disease. The present	
CC	sequence is that of the Gene 216 genomic nucleic acid sequence, promoter	
CC	or enhancer	
XX		
XX		
SQ	Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;	
	Query Match	2.5%; Score 35; DB 5; Length 207433;
	Best Local Similarity	55.3%; Pred. No. 92;
	Matches	68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY	1210 GGGCTACGTAGTACATAGAGTCATAGAACTATAAGAGCTTAAGACCATTCACAT	1266
DB	1335 GTCCACCTCTTCTCTTAAACACAAATACCAATTCCTCCAGGGCTATAGACCTCTTGAAT	1276
QY	1270 CCTAAGGGCTCTCTCTTCTTTCGCATCATCATGAATCATCTCACTCAACAGGAACTC	1328
DB	1275 ATACGAACATCTCTTCTCTTAAGGAAACCTCAGGATATATTAATCCAAATAGAAACC	1316
QY	1330 TTT 1332	
DB	1215 ATT 1213	
RESULT 13		
ID	ABX74891/C	
AC	ABX74891 standard; DNA; 207433 BP.	
XX	ABX74891;	
XX		
DT	07-APR-2003 (first entry)	
XX		
DE	BAC1098L22 DNA sequence.	
XX		
KW	Gene 216; antiasthmatic; antiinflammatory; ss; anorectic;	
KW	chromosome 20p13-p12; single nucleotide polymorphism; SNP; gene therapy;	
KW	respiratory disease; asthma; obesity; bronchial hyper-responsiveness;	
KW	chronic obstructive pulmonary disease;	

XX	adult respiratory distress syndrome; inflammatory bowel syndrome.	XX
XX	Synthetic.	XX
XX	WO200283077-A2.	XX
XX	24-OCT-2002.	XX
XX	15-APR-2002; 2002WO-US012063.	XX
XX	13-APR-2001; 2001US-00834597.	XX
XX	13-APR-2001; 2001WO-US012245.	XX
XX	(SCHE) SCHERING CORP.	XX
XX	(GENO-) GENOME THERAPEUTICS CORP.	XX
XX	Keith T, Little RD, Van Berdeghen P, Dupuis J, Del Mastro RG;	XX
XX	Simon J, Allen K, Pandit S;	XX
XX	WPI; 2003-092960/08.	XX
XX	New isolated gene 216 nucleic acids, useful for diagnosing, preventing or	XX
XX	treating a disorder, such as asthma, bronchial hyper-responsiveness,	XX
XX	chronic obstructive pulmonary disease, obesity or inflammatory bowel	XX
XX	syndrome.	XX
XX	Example 6; Fig 7; 650bp; English.	XX
XX	This invention relates to a novel isolated nucleic acid, gene 216,	XX
XX	identified from human chromosome 20p13-p12. The invention also discloses	XX
XX	regions of the 216 gene that contain single nucleotide polymorphisms	XX
XX	(SNP's) which may be used as markers for disease susceptibility or	XX
XX	severity. The nucleotides of the invention may have antiasthmatic,	XX
XX	antiinflammatory or anorectic activities and may be used in gene therapy.	XX
XX	The nucleic acids, antibodies or its fragments are useful for diagnosing,	XX
XX	preventing or treating a disorder, such as respiratory diseases (e.g.	XX
XX	asthma, bronchial hyper-responsiveness, chronic obstructive pulmonary	XX
XX	disease or adult respiratory distress syndrome), obesity, or inflammatory	XX
XX	bowel syndrome. The nucleic acids are also useful for identifying	XX
XX	increased susceptibility of a subject to the disorders mentioned. The	XX
XX	nucleic acids can also be used as primers and templates for the	XX
XX	recombinant production of disorder-associated peptides or polypeptides,	XX
XX	for chromosome and gene mapping, or for tissue distribution studies. The	XX
XX	present sequence represents a gene 216 cDNA sequence used in the scope of	XX
XX	the invention	XX
XX	Sequence 207433 BP; 52775 A; 51289 C; 51698 G; 51671 T; 0 U; 0 Other;	XX
XX	Query Match 2.5%; Score 35; DB 8; Length 207433;	XX
XX	Best Local Similarity 55.3%; Pred. No. 92;	XX
XX	Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;	XX
XX	1210 GCGGTAGCTAGATACATAGAGTCTATAGAAAATATAAGAGCTTGAAACCATTCAAAT 1269	XX
XX	1335 GTCTTACCTTCTTTTAAACACAAATACCAAAATCCAGGGCTGTGGAGCCCTTGAAT 1276	XX
XX	1270 CCTAAGGCTCTCTCTTCTTCTTGATCATCATCAAGATCATCACTCAACACGAACTC 1329	XX
XX	1215 ATACGAAACATCTCTTCTTAAGGAAAACCTCAGGATTAATTAATCCAAATAGAAACC 1216	XX
XX	1330 TTT 1332	XX
XX	1215 ATT 1213	XX
XX	Db	XX
XX	RESULT 14	XX
XX	ADJ36614/C	XX
XX	ID ADJ36614 standard; DNA; 207433 BP.	XX
XX	ADJ36614;	XX
XX	DT 22-APR-2004 (first entry)	XX

DE Bacterial artificial chromosome RBC1-11.
 XX antiasthmatic; respiratory; gene therapy; asthma;
 KW bronchial hyperresponsiveness; atopy; chronic obstructive lung disease;
 KW adult respiratory distress syndrome; obesity; inflammatory bowel disease;
 KW human; gene 216; bacterial artificial chromosome; BAC; RBC1-11; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 US2004002470-A1.
 XX
 PD 01-JAN-2004.
 XX
 PF 17-OCT-2002; 2002US-00277216.
 XX
 PR 13-APR-2000; 2000US-00548797.
 PR 13-APR-2001; 2001US-00834597.
 PR 19-APR-2002; 2002US-00126022.
 XX
 PA (KEIT/) KEITH T.
 PA (LITT/) LITTLE R D.
 PA (VEER/) VAN EERDEWEGH P.
 PA (DUPU/) DUPUIS J.
 PA (DMAS/) DEL MASTRO R G.
 PA (SIMO/) SIMON J.
 PA (ALLE/) ALLEN K.
 PA (PAND/) PANDIT S.
 XX
 PI Keith T, Little RD, Berdewegh PV, Dupuis J, Del Mastro RG;
 PI Simon J, Allen K, Pandit S;
 XX
 DR WPI; 2004-061675/06.
 XX
 PT Gene 216 nucleic acid, useful for preparing a composition for treating
 PT disorders e.g., asthma, bronchial hyperresponsiveness, atopy, chronic
 PT obstructive lung disease and adult respiratory distress syndrome.
 XX
 PS Example 6; SEQ ID NO 5; 441bp; English.
 XX
 CC The invention describes a new isolated nucleic acid comprising a fully
 CC defined sequence having 23574 bp or at least its 50 or 15 contiguous
 CC nucleotides and includes: allele G of single nucleotide polymorphism
 CC (SNP) AB+2; allele G of SNP BC+1; and allele C of SNP BC+2. The invention
 CC describes identifying increased susceptibility to a disorder comprising
 CC asthma, bronchial hyperresponsiveness, atopy, chronic obstructive lung
 CC disease and adult respiratory distress syndrome in a subject comprising
 CC testing a biological sample obtained from a subject for the presence of
 CC at least one allele or haplotype given in the specification, where the
 CC presence identifies an increased susceptibility to the disorder. The
 CC nucleic acid is useful for preparing a composition for treating disorders
 CC comprising asthma, bronchial hyperresponsiveness, atopy, chronic
 CC obstructive lung disease and adult respiratory distress syndrome. This
 CC sequence represents a bacterial artificial chromosome (BAC) containing
 CC human gene 216 isolated from chromosome 20p13-p12 associated with asthma,
 CC obesity and inflammatory bowel disease.
 CC
 SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
 XX
 Query Match 2.5%; Score 35; DB 12; Length 207433;
 Best Local Similarity 55.3%; Pred. No. 92;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 1210 GGCCTACGTAGATGATGAGTAAAGATTAAGAGCTTGAAGACCATTCAT 1269
 DB 1335 GTCTTACCTCTTCTTAAACACATTAACAAATTCCTCCAGGCTAGTGAACCTTGAAT 1276
 QY 1270 CCTAAGGCTCTCTTCTTCTTGTGATCATCATCAAGATTCATCACTCAACAGGAATCTC 1329
 DB 1275 ATACGAAACATCTCTTCTTCTTAAAGAAACCTCAGGATTAATTAATTCCTCAATTAAGAAACCC 1216
 QY 1330 TTT 1332
 ||

DB 1215 ATT 1213
 RESULT 15
 ADL81193/c
 ID ADL81193 standard; DNA; 207433 BP.
 XX
 AC ADL81193;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE BAC1098L22 DNA sequence.
 XX
 KW asthma; bronchial hyperresponsiveness; obesity;
 KW inflammatory bowel disease; human; ds; gene 216.
 XX
 OS Homo sapiens.
 OS
 XX
 US2004023215-A1.
 XX
 PD 05-FEB-2004.
 XX
 PF 19-APR-2002; 2002US-00126022.
 XX
 PR 13-APR-1999; 99US-0129391P.
 PR 13-APR-2000; 2000US-00548797.
 PR 13-APR-2001; 2001US-00834597.
 XX
 PA (KEIT/) KEITH T.
 PA (LITT/) LITTLE R D.
 PA (BERD/) BERDEWEGH P V.
 PA (DUPU/) DUPUIS J.
 PA (DMAS/) DEL MASTRO R G.
 PA (SIMO/) SIMON J.
 PA (ALLE/) ALLEN K.
 PA (PAND/) PANDIT S.
 XX
 PI Keith T, Little RD, Berdewegh PV, Dupuis J, Del Mastro RG;
 PI Simon J, Allen K, Pandit S;
 XX
 DR WPI; 2004-142647/14.
 XX
 PT New isolated nucleic acid molecules useful for diagnosing or treating
 PT asthma or bronchial hyperresponsiveness, or other diseases such as
 PT obesity or inflammatory bowel disease.
 XX
 PS Example 5; SEQ ID NO 5; 485bp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule, or a set of
 CC nucleic acid molecules each given in the specification. The composition
 CC and methods are useful in diagnosing or treating asthma or bronchial
 CC hyperresponsiveness, and other diseases such as obesity or inflammatory
 CC bowel disease. The present sequence is used in the exemplification of the
 CC present invention.
 CC
 SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
 XX
 Query Match 2.5%; Score 35; DB 12; Length 207433;
 Best Local Similarity 55.3%; Pred. No. 92;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 1210 GGCCTACGTAGATGATGAGTAAAGATTAAGAGCTTGAAGACCATTCAT 1269
 DB 1335 GTCTTACCTCTTCTTAAACACATTAACAAATTCCTCCAGGCTAGTGAACCTTGAAT 1276
 QY 1270 CCTAAGGCTCTCTTCTTCTTGTGATCATCATCAAGATTCATCACTCAACAGGAATCTC 1329
 DB 1275 ATACGAAACATCTCTTCTTCTTAAAGAAACCTCAGGATTAATTAATTCCTCAATTAAGAAACCC 1216
 QY 1330 TTT 1332
 DB 1215 ATT 1213

, Fri Feb 25 18:19:08 2005

us-10-070-386-1.rng

Page 11

Search completed: February 9, 2005, 10:48:33
Job time : 757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:24:55 ; Search time 272 Seconds
(without alignments)
8289.667 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtcacgctgggtgctgatat.....accatcatcaactaca 1378

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUG_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	42.2	3.1	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	37.6	2.7	289	3 US-09-007-005-17	Sequence 17, Appl
C 3	37.6	2.7	289	3 US-09-244-796-17	Sequence 17, Appl
C 4	35.2	2.6	399	4 US-09-621-976-8976	Sequence 8976, Ap
C 5	35.2	2.6	505	4 US-09-621-976-15639	Sequence 15639, A
C 6	35.2	2.5	136917	4 US-09-949-016-16369	Sequence 16369, A
C 7	34.4	2.5	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 8	34.4	2.5	16871	4 US-09-949-016-12718	Sequence 12718, A
C 9	34.4	2.5	16874	4 US-09-949-016-15176	Sequence 15176, A
C 10	34.2	2.5	1500	4 US-09-107-532A-1870	Sequence 1870, Ap
C 11	34.2	2.5	134890	4 US-09-949-016-15602	Sequence 15602, A
C 12	33.4	2.4	4792	3 US-08-781-891-205	Sequence 205, App
C 13	33.4	2.4	4792	3 US-09-618-166-205	Sequence 205, App
C 14	33.4	2.4	6476	3 US-09-127-670-5	Sequence 208, App
C 15	33.4	2.4	16442	3 US-08-781-891-208	Sequence 208, App
C 16	33.4	2.4	16442	4 US-09-618-166-308	Sequence 208, App
C 17	33.3	2.4	364	4 US-09-621-976-17202	Sequence 17202, A
C 18	32.8	2.4	414	4 US-09-328-352-1277	Sequence 1277, Ap
C 19	32.8	2.4	414	4 US-09-806-708B-22	Sequence 22, Appl
C 20	32.8	2.4	1416	4 US-09-902-540-872	Sequence 872, App
C 21	32.6	2.4	10813	4 US-08-956-171B-364	Sequence 364, App
C 22	32.6	2.4	10813	4 US-08-781-986A-364	Sequence 364, App
C 23	32.4	2.4	630	4 US-09-248-786A-5221	Sequence 5221, Ap
C 24	32.4	2.4	28960	4 US-09-949-016-11978	Sequence 11978, A
C 25	32.4	2.4	30054	4 US-09-949-016-16100	Sequence 16100, A
C 26	32.2	2.3	978	3 US-09-267-031-13	Sequence 13, Appl
C 27	32.2	2.3	89843	4 US-09-949-016-12346	Sequence 12346, A

28	32.2	2.3	89844	4 US-09-949-016-13656	Sequence 13656, A
29	32	2.3	988	3 US-08-793-634B-7	Sequence 7, Appl
30	32	2.3	1281	3 US-09-359-070-1	Sequence 1, Appl
31	32	2.3	1300	2 US-08-440-845D-8	Sequence 8, Appl
32	32	2.3	1300	3 US-08-868-458-8	Sequence 8, Appl
33	32	2.3	1300	3 US-09-134-262-1	Sequence 1, Appl
C 34	32	2.3	10445	4 US-09-949-016-12311	Sequence 12311, A
C 35	32	2.3	10445	4 US-09-949-016-12927	Sequence 12927, A
C 36	32	2.3	42741	4 US-09-949-016-11857	Sequence 11857, A
C 37	31.8	2.3	69874	4 US-09-949-016-12361	Sequence 12361, A
38	31.8	2.3	69874	4 US-09-949-016-13049	Sequence 13049, A
39	31.8	2.3	94077	4 US-09-949-016-13635	Sequence 13635, A
C 40	31.8	2.3	278866	4 US-09-949-016-13922	Sequence 13922, A
C 41	31.8	2.3	278866	4 US-09-949-016-13923	Sequence 13923, A
C 42	31.8	2.3	278866	4 US-09-949-016-13924	Sequence 13924, A
C 43	31.8	2.3	278866	4 US-09-949-016-13925	Sequence 13925, A
C 44	31.8	2.3	278866	4 US-09-949-016-13926	Sequence 13926, A
C 45	31.8	2.3	278866	4 US-09-949-016-14699	Sequence 14699, A

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-F18
US-08-232-463-14
Query Match 3.1%, Score 42.2, DB 1, Length 7218,

```

RESULT 2
US-09-007-005-17
; Sequence 17, Application US/0907005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szoatak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihue
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

```

Db      1  RGRGRGARCARARURURARCBURARURURUDARCBARARURURARCBARARURGRNRN 60
Qy      1095  TGGCTCTTCGCGCTGTCACTTACACACATCATGGGGCTCAACACGATACAGCTTCA 1154
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 120
Qy      1155  AGAGTGGCGCATGGAAGTACTCCGATCGAACC GGAAAGCGGTTCAACATCGGCGT 1214
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 180
Qy      1215  ACCTGATACATAGAGTCATAGAAACATAAAGAGCCTTGAGAAC 1260
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRCRAGRC 226

RESULT 3
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szoetak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      2.7%; Score 37.6; DB 3; Length 289;
Best Local Similarity 4.0%; Pred. No. 0.035;
Matches 9; Conservative 103; Mismatches 114; Indels 0; Gaps 0;

Qy      1035  GGGCGAGAGCAACCTGATTTTCAATGAATCAATCAATTCGCGGAGAGTTCCAA 1094
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      1  RGRGRGARCARARURURARCBURARURURUDARCBARARURURARCBARARURGRNRN 60
Qy      1095  TGGCTCTTCGCGCTGTCACTTACACACATCATGGGGCTCAACACGATACAGCTTCA 1154
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 120
Qy      1155  AGAGTGGCGCATGGAAGTACTCCGATCGAACC GGAAAGCGGTTCAACATCGGCGT 1214
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 180
Qy      1215  ACCTGATACATAGAGTCATAGAAACATAAAGAGCCTTGAGAAC 1260
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRCRAGRC 226

RESULT 4
US-09-621-976-8976/C
; Sequence 8976, Application US/09621976
; Patent No. 6639063

```


Query Match 2.5%; Score 34.4; DB 4; Length 832;
Best Local Similarity 10.4%; Pred. No. 0.93;
Matches 34; Conservative 155; Mismatches 136; Indels 1; Gaps 1;

```
OY 933 TAAGCTCTGGAAATACAACTGTGATTAACCAATTCATCGTCAAGGTCGCTGATATG 992
DB 46 KAMWKYMTWTWMMWYMMGTYKKKAMCRKTGKKKKKGYMMWYMMRBSYAMMMWTRT 105
OY 993 CTGAGCTTCCCTGCGATGTGGGTAACTATTTCATGAGGGCGAGATGCAACTCT 1052
DB 106 WTGAAYYSMMYMRCKKKAYRKTTCYSSKMTWMMKKAATTTMMKKTYMAAT 165
OY 1053 ATTTCAATTGAATCTAACTATTCTGGGTAGAGAGATTCTCAATGCTTCTCGCTGAC 1112
DB 166 RYMMWMCWTKEMRASMMYCMWMMKARKSWRKSRYSASBARCKCYSCSMGAMSMYKM 225
OY 1113 TTACACACATCATGGGGGTCAACAACGTATACAGTTCTATAGAGAGTCGGCATTAAGT 1172
DB 226 WRMRMRGMATGAGKAWRASCMWRKRYAGSKTSYKSMWMCWTRSMY -CYTKARMTGY 284
OY 1173 AGCTACCGCATCGAACCAGGCTTCAAGACATGGCGGTAGCATATCATAGAGTGC 1232
DB 285 YCYKKGMMGKRGKRYASKTKTWMMKMMWCMARMTYISTGTASMMWRMTYTTMMKMMKY 344
OY 1233 ATAGAACATTAAGAGACCTTGAAAGA 1258
DB 345 AWAABAAMRWMMWMMWMBRACAAATA 370
```

RESULT 8

US-09-949-016-12718
; Sequence 12718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12718
; LENGTH: 16871
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12718

Query Match 2.5%; Score 34.4; DB 4; Length 16871;
Best Local Similarity 52.0%; Pred. No. 7.4;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
OY 875 TGGCGAAAACAGTTTAACTTAAGTAAAGGAGAGCTGAACCTGTGGAATA 934
DB 489 TGGGAGAGACTGGGGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
OY 935 AGCCTGTGGAATCAACAGTTGATTAACCAATTCAGTCAAGGTCGCTGATATGCT 994
DB 549 TTAAGTATTTTTCAAAGTACTTTAAAAATGTAATCTTTATCTATATGATATGCA 608
OY 995 GGAGCTTCCCTGCTGCATTTGGGGTAA 1022
DB 609 AAATATTCATGACACCTTGTGAGTGA 636
```

RESULT 9

US-09-949-016-15176
; Sequence 15176, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15176
; LENGTH: 16874
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15176

Query Match 2.5%; Score 34.4; DB 4; Length 16874;
Best Local Similarity 52.0%; Pred. No. 7.4;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
OY 875 TGGCGAAAACAGTTTAACTTAAGTAAAGGAGAGAGAGAGAGAGAGAGAGAG 934
DB 489 TGGGAGAGACTGGGGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
OY 935 AGCCTGTGGAATCAACAGTTGATTAACCAATTCAGTCAAGGTCGCTGATATGCT 994
DB 549 TTAAGTATTTTTCAAAGTACTTTAAAAATGTAATCTTTATCTATATGATATGCA 608
OY 995 GGAGCTTCCCTGCTGCATTTGGGGTAA 1022
DB 609 AAATATTCATGACACCTTGTGAGTGA 636
```

RESULT 10

US-09-107-532A-1870/C
; Sequence 1870, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489

```

1 REFERENCE/DOCKET NUMBER:  GTC-012
2 TELECOMMUNICATION INFORMATION:
3 TELEPHONE: (781)893-5007
4 TELEFAX: (781)893-8277
5 INFORMATION FOR SEQ ID NO: 1870:
6
7 SEQUENCE CHARACTERISTICS:
8
9 LENGTH: 1500 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: double
12 TOPOLOGY: circular
13 MOLECULE TYPE: DNA (genomic)
14 HYPOTHETICAL: NO
15 ANTI-SENSE: NO
16 ORIGINAL SOURCE:
17 ORGANISM: Enterococcus faecium
18
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (b) LOCATION 1..1500
22 SEQUENCE DESCRIPTION: SEQ ID NO: 1870
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1
```

Query Match	2.5%	Score 34.2;	DB 4;	Length 1500;
Best Local Similarity	49.2%	Pred. No. 1.6;		
Matches 90; Conservative	0;	Mismatches 93;	Indels 0;	Gaps 0;

Qy	1042	AATGCACTCAATTTTCATTTGAATCTAAATCTAACTATCTGGTAGAGAGTTCTCAATAGTCTT	1101
Db	1219	AATGAAGCTCTTGTAGAGATGCTTCCACATCAATTTGTGTGTGTATTTCCAGATACACTTT	1160
Qy	1102	CTCGCTGTCACCTTACACATCATATGGGGGTCAACAACGTATACAGCTTCATAGAGAGTGC	1161
Db	1159	CATAATTTTGAATCCATATTTTCAGTAGTTACGAAACATCTTACCCCTTCATTTGAAAAATCA	1100
Qy	1162	GGCATTTGAAGTAGCTACCGCATTCGAAACCCGGAAGCGTTCAAGACATGGGCGTAGTGA	1221
Db	1099	TTTTTTTCATTTTCATTAACGAATGATATATGAGAGGGTAAATGCTGTCTGCTCGA	1040
Qy	1222	TAC 1224	
Db	1039	TAC 1037	

RESULT 11
US-09-949-016-15602
; Sequence 15602, Application US/09949016

```

? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO: 15602
? LENGTH: 134890
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-15602

```

Query Match	2.5%	Score 34.2;	DB 4;	Length 134890;
Best Local Similarity	56.8%	Pred. No. 36;		
Matches 63;	Conservative 0;	Mismatches 48;	Indels 0;	Gaps 0;

1155 AGAGTGGCGGCTTGAAGTAGCTACCGCATGGAACCGGAAAGCGGTCAAGACATGGGGCGT 1214

Db 102216 AAGTGAAGGTTCAAGCATATCGAATCTCACTCTGAGTGTGGAACAAAGTCT 1022175

Qy 1215 ACGTAGATCATAGAGTCATAGAAACATTAAGAAGCTTTGAAGAACCATTC 1265

Db 102276 CAGTGATTTAAAGGTTCAAGSTAACTAAATGTTTAGATGAAGACATATTC 102326

RESULT 12
US-08-781-891-205

GENERAL INFORMATION:
APPLICANT: Fu, Ying-hui
APPLICANT: Yu, Chang-hn
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;

```

APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: NO. 6090620tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 205:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4792 base pairs
;

```

```
;;  
; type: nucleic acid  
; strandedness: single
```

```

;      TOPOLOGY: linear
;
;      FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 145..4347
;

```

US-08-781-891-205

Query Match	Score	DB	Length
2.4%	33.4	3	4792

Best Local Similarity 49.7%; Pred. NO. 6.8;
Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0.

213 ATGACACCCACCCCGCAACGGAACTCTGAGTTCAAGAGGGTGTGCTTACGGCATTTTA 272

Db 1164 ATGGGACCCAGACTTGACAGTTTAGTGAAGCAAGAGGAGGTGATGTATTAGAAATCA 1222

273 GGTATAGATGCATAGGGTTGACGTACGTGAAGCTGATTACGAGACATGAGACAAAG 332

Db
1224 AGTGAGCGAAGAAAAAGTGAACTCTGAAATGAAATGAGAGACAATCTGTTGAGAGAGA 1283

333 AAATACACGGTTGTATGCGTTCCGCTTACTTAAAGTATATCCACA 383

Db 1284 TATGGAAGAACTGTGTGATTCTAGTATTTCAGAAATGAACTCCAGA 1334

RESULT 13
US-09-618-166-205

RESULT 13
US-09-618-166-205

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-208

Query Match 2.4%; Score 33.4; DB 3; length 16442;
Best Local Similarity 49.7%; Pred. No. 16;
Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

OY 213 ATGACACCCACCCCGCAACGGAATCCTGAGTTCAAGAGGGGTGCTCTACGGCATTTA 272
13744 ATGGACCCAGAACTTGACAGTTTAGTAGCAAGAGGGTTGATGTATTTAGAAATCA 13803
OY 273 GGTATAGATGCGCATAGGGTTTGACGTAAGCTGAAAGCTGATTACGACATGAGACAAAG 332
13804 AGTGAAGCAAGAAAAGTGAATCTGAAAATGAAATAGAAAGATAATCTGTTGAGAGAAAG 13863
OY 333 AAAATACAAAGGTGTATGCGTTCCCGTCTTACTAAAGTGATATCCAAGA 383
13864 TATGGAAGAAGCTGTGTGATTCCTAGTATTCAGAAAATGAATCCAAGA 13914
DB

Search completed: February 9, 2005, 13:44:38
Job time : 275 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 9, 2005, 12:28:25 ; Search time 763 Seconds
(without alignments)
10395.036 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtcgacgcygggtgtgatcat.....accatcatcaactaaca 1378

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.4	2.6	600	US-10-123-155-462	Sequence 462, App
2	36.4	2.6	600	US-10-146-731-462	Sequence 462, App
3	36.4	2.6	600	US-10-140-472-462	Sequence 462, App
4	36.4	2.6	600	US-10-141-761-462	Sequence 462, App
5	36.4	2.6	600	US-10-142-885-462	Sequence 462, App
6	36.4	2.6	600	US-10-158-790-462	Sequence 462, App
7	36.4	2.6	600	US-10-137-871-462	Sequence 462, App
8	36.4	2.6	600	US-10-140-923-462	Sequence 462, App
9	36.4	2.6	600	US-10-141-756-462	Sequence 462, App
10	36.4	2.6	600	US-10-141-759-462	Sequence 462, App
11	36.4	2.6	600	US-10-140-805-462	Sequence 462, App

12	36.4	2.6	600	US-10-140-864-462	Sequence 462, App
13	36.4	2.6	600	US-10-142-426-462	Sequence 462, App
14	36.2	2.6	2672	US-09-775-938A-34	Sequence 34, App1
15	36	2.6	3673778	US-10-312-841-1	Sequence 1, App1
16	35.8	2.6	134738	US-10-087-192-502	Sequence 502, App
17	35.4	2.6	993	US-09-898-837A-8	Sequence 8, App1
18	35	2.5	458	US-10-184-644-198	Sequence 458, App
19	35	2.5	458	US-10-184-634-198	Sequence 458, App
20	35	2.5	537	US-10-027-632-774500	Sequence 274500,
21	35	2.5	537	US-10-027-632-774500	Sequence 274500,
22	35	2.5	6971	US-10-311-455-1210	Sequence 1210, App
23	35	2.5	207433	US-10-277-216-5	Sequence 5, App11
24	35	2.5	207433	US-10-126-022-5	Sequence 36862, A
25	34.8	2.5	4416	US-10-369-493-16862	Sequence 98, App1
26	34.4	2.5	8805	US-10-257-166-98	Sequence 154, App
27	34.4	2.5	32874	US-10-052-482-154	Sequence 7196, App
28	34.2	2.5	312	US-10-674-124A-7196	Sequence 160, App
29	34.2	2.5	90435	US-10-332-696-160	Sequence 39026, A
30	34	2.5	1591	US-10-437-963-39026	Sequence 807, App1
31	34	2.5	6283	US-10-311-455-807	Sequence 73, App1
32	34	2.5	6283	US-10-257-166-73	Sequence 16934, App1
33	34	2.5	339234	US-10-322-696-73	Sequence 16934,
34	33.8	2.5	818	US-10-027-632-166934	Sequence 166934,
35	33.8	2.5	818	US-10-027-632-166934	Sequence 166934,
36	33.8	2.5	818	US-10-027-632-166934	Sequence 166934,
37	33.8	2.5	818	US-10-027-632-166934	Sequence 166934,
38	33.6	2.4	25179	US-10-322-681-277	Sequence 277, App
39	33.4	2.4	403	US-09-864-761-5003	Sequence 5003, App
40	33.4	2.4	737	US-10-425-115-65924	Sequence 65924, A
41	33.4	2.4	1024	US-10-123-155-198	Sequence 198, App
42	33.4	2.4	1024	US-10-146-731-198	Sequence 198, App
43	33.4	2.4	1024	US-10-140-472-198	Sequence 198, App
44	33.4	2.4	1024	US-10-141-761-198	Sequence 198, App
45	33.4	2.4	1024	US-10-142-885-198	Sequence 198, App

ALIGNMENTS

RESULT 1
US-10-123-155-462
; Sequence 462, Application US/10123155
; Publication No. US20030068794A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-462

```

Query Match      2.6%; Score 36.4; DB 14; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGATGTGAGACATTTTCTAAGCATCCATA 667
DB 1 MRSCIMRCRHLISQGVQMSLLAVLVPFLPALPSFIKEPQTPRSRHOETENIKERLSGLA 60
QY 668 TAGGATTTGGCGGCGTAATGCGGATTTGAAGGATTAAGGGGGTGAAGTGTGTCTC 727
DB 61 KPKSQAPFRARRTTIYAEPAENNALNTOTQPKATTTGDRKEANQAPPEEQDKVPHNQ 120
QY 728 AAAAGAGGTGATGCTATATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTCTGACG 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWSQDTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATPAAGCTCACTTGTAGTATGTCACCTTAATGTGACGAGATGCAATGCTGAT 847
DB 181 SEKHQKAAATTAATKTLIPKSHMLAPTGAVSTRTRQKGVTTAVIPPKKKPQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAACAGTTAGATCTAGTTAAAGGA 907
DB 241 QSPTRORNRUKANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKASKSLMLQKLEPN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAAATGAAGCTGTTGAATACACGTTGATTAACCAATT 967
DB 301 LTLFLDSRHFQSEWDRLEHFAFPFGFMELNYSLVQKVTTRPPVPOQULLASLPAGSL 360
QY 968 CAGTGTCAAGGTCCTCGTATATGCTGAGCTTCCCTGCTGATTTGGGTAATCTATT 1027
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYVFLSGALIKYBQDVGRTRSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTAATGAATCTAATCTATTTCTGGTAGAG 1087
DB 421 SLLILGNRQFKNVPKQDVRYLHFLBGRDYEMLEALLMNQTVMSKNLFWFRRHQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 2
US-10-146-731-462
; Sequence 462, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600

```

```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-462

Query Match      2.6%; Score 36.4; DB 15; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGATGTGAGACATTTTCTAAGCATCCATA 667
DB 1 MRSCIMRCRHLISQGVQMSLLAVLVPFLPALPSFIKEPQTPRSRHOETENIKERLSGLA 60
QY 668 TAGGATTTGGCGGCGTAATGCGGATTTGAAGGATTAAGGGGGTGAAGTGTGTCTC 727
DB 61 KPKSQAPFRARRTTIYAEPAENNALNTOTQPKATTTGDRKEANQAPPEEQDKVPHNQ 120
QY 728 AAAAGAGGTGATGCTATATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTCTGACG 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWSQDTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATPAAGCTCACTTGTAGTATGTCACCTTAATGTGACGAGATGCAATGCTGAT 847
DB 181 SEKHQKAAATTAATKTLIPKSHMLAPTGAVSTRTRQKGVTTAVIPPKKKPQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAACAGCTTGAATCTAGTTAAAGGA 907
DB 241 QSPTRORNRUKANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKASKSLMLQKLEPN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAAATGAAGCTGTTGAATACACGTTGATTAACCAATT 967
DB 301 LTLFLDSRHFQSEWDRLEHFAFPFGFMELNYSLVQKVTTRPPVPOQULLASLPAGSL 360
QY 968 CAGTGTCAAGGTCCTCGTATATGCTGAGCTTCCCTGCTGATTTGGGTAATCTATT 1027
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYVFLSGALIKYBQDVGRTRSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTAATGAATCTAATCTATTTCTGGTAGAG 1087
DB 421 SLLILGNRQFKNVPKQDVRYLHFLBGRDYEMLEALLMNQTVMSKNLFWFRRHQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 3
US-10-140-472-462
; Sequence 462, Application US/10140472
; Publication No. US20030138888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm

```

```

; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-462

```

Query Match 2.6%; Score 36.4; DB 15; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

Qy	608	AGACCTACATATGCGCATCTTAAAGCAGGATCTGAGACATTTCTTAAGGCATCCATA	667
Db	1	MRSLCMRKHLSQGVQWMLLAVLVFLPALPSFIKEPOTKPSKHORTIENTKERSLOSLA	60
Qy	668	TAGGACATTGGCGCGCTAAAGTCGGCATTTGAAGAGATAAGGGGGTGTGAAGTGTGTCTC	727
Db	61	KPKSGAPTRARRTTYAEPARPENNALNTOTQPKAHTTDDRGKMANQAPREBDKVPHTAQ	120
Qy	728	AAAAGAGAGTGGATTGGCTATTACACGCGCTAAGCAGGTGGGCTAGCAGACTGTCTGCAGC	787
Db	121	RAAWKSPEKEKTMVNTLSPROGDAMAGRETAQSKESQDTKTTQGNNGQUTKXLTASRTV	180
Qy	788	TGTGAATTAAGCTCACTTGTCTAAGTATGTCCACCTTAAGTATGTACGAGAGTGAATTCGTAAT	847
Db	181	SEKIOGKATTAKTAKTLIPKSOHNMILAPTAVSTRTRQKGVTTAVIPRKEKKPDPATPPAPF	240
Qy	848	TGGGTTAAATAGGGCATGATGATGATGATGCGGCAAAAACGTTTAATCATCTAATTAAAGGA	907
Db	241	QSPPTTORNRQLKANFKSEBPRIIDFEKKSFEIGLQTTCPBDSVKIKASISLYLQKLFLEN	300
Qy	908	AGCTGAAGCTGAACCTGTCTGCAAAATAAGCCTGTGGAATACAAAGTTGATNAACCAATT	967
Db	301	LTLFLDSHNFGQSEMDRLHNHFAFRPFGMYELNYSLVQKVVTRPRPRQOQLLASLPRASL	360
Qy	968	CAGTCGTCAAGGGTGTCTGTATATCTGAGAGCTTCCCTGTGCGATTGTGGGGTAACATAAT	1027
Db	361	RCITCAVVGNSGILINNSHQEVIDSHDVYFRLSGALIKGYEODVGRTRISFYGTAFSLTQ	420
Qy	1028	TCATAGTGGGCGAATGCAACTATTTCTCAATGAATCTPAACATAATTCGTGGGTAGGAG	1087
Db	421	SLLLIGNNGFNQVPLGKDVRYLHLFDEGTRDYEMLBALMNQTMVSKNLFWFHRHPOEAFR	480
Qy	1088	TTTCAATGATGTCTT	1101
Db	481	EALHMDRYLLHPD	494

```

RESULT 4
US-10-141-761-462
; Sequence 462, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Beresini, Kevin P.
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tuma, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198

```

```

CURRENT APPLICATION NUMBER: US/10/141,761
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO. 462
LENGTH: 600
TYPE: PR1
ORGANISM: Homo Sapien
US-10-141-761-462

```

Query Match	2.6%;	Score 36.4;	DB 15;	Length 600;
Best Local Similarity	9.7%;	Pred. No. 1.5;		
Matches 48;	Conservative 125;	Mismatches 321;	Indels 0;	Gaps 0;

Qy	608	AGACCTTCATCATCCCACTGTTAAAGACGGAGTCTGACATTTTCTAAAGACCATTA	667
Db	1	MRSCLMCRHLSGVQVMSLLAVLVFLFALPSPFIKEQPTSRHQRTENIKERLSGLA	60
Qy	668	TAGCATTGGCGCGCTAAGTCGCATTGAAGAGATAAGGGGGGTGTGAAGTGTGTC	727
Db	61	KPKSGAFTBARRTTIYAEPAPENNALNTQTPKXHTTGDRCKANQAPREBDQKVPHTAQ	120
Qy	728	AAAAGAGAGTCGATTGGCTATACACGCCGCTAAGCAGGTGGCTAGCAGCTGTCTGCAGC	787
Db	121	RAAKSPSEKEKTMVNTLSPRGODAGMASGRTEAOSWKSQDPTKTTQSGNGQTRKLLASRTV	180
Qy	788	TGTGAATTAAGCTACTTGGCTTAGTAGTATGCCACTTAATGTACAGCAGATGCAATTCGTAT	847
Db	181	SEKIOGKAATTAKTLLIPKSOHRMLAPGVASTRORQVTAIVLPRPEKKQAPTRPAPE	240
Qy	848	TGGGTTAAATGGGCGATGATAGTAGAGGCGCGAAACAAGCTTTAGATCTAGTTAAAGGA	907
Db	241	QSPFTQNRQRLKANKFSPBRWDDEKKSFETGGIQTTCPPSVKIKASKSLMLQTLFIPN	300
Qy	908	AGCTGAAGCTGAACCTGTCAGAATAGACCTGTGGAAATCAACGTTGATTAACCAATT	967
Db	301	LTLFLDSRHFNQSWMDLHNHAPRFGFHEMLNYSLVQKVTFRPPRPQOQLLASLPASLS	360
Qy	968	CAGTCGTCAAGGCTGCTCTGATATGCTGGAGCTTCCCTGCGCATTTGGGGTAACTATT	1027
Db	361	RCITCAVUNGNGILNNSHMQEIDSHDVFRLSGALIKGYEODVGRTRSPFGTAFSLTQ	420
Qy	1028	TCATATGGGGGAGAGATGCAACTATTTTCAATTGAATCTTAACATAATTCGGGTAGGAG	1087
Db	421	SLTLIGNRGFKNVLBGKDVRYLHFLBGRDYEMLEALLMOTVMSKDLFWFRHBPQEAER	480
Qy	1088	TTTCATATGGCTTT 1101	
Db	481	EALHMDRYLLHHPD 494	

```

RESULT 5
US-10-142-085-462
/ Sequence 462, Application US/10142885
/ Publication No. US20030157604A1
/ GENERAL INFORMATION:
/
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Denoyer, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godwami, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tunas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zhenlin

```

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P330R1C248
;; CURRENT APPLICATION NUMBER: US/10/142,885
;; CURRENT FILING DATE: 2002-05-10
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 462
;; LENGTH: 600
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-142-885-462

Query Match 2.6%; Score 36.4; DB 16; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGGATCTGAGACATTTCTTAAGGCATCCATA 667
DB 1 MRSCLMCRHLISQGVMSLLAVLVFLFALPSPFIKEPQTPSRHQRTEIKERSLSLSLA 60
QY 668 TAGGCATTGGCGCGTAAGTCCGATTTGAAGGATTAAGGGGGGTGTAAGTGTGTCTC 727
DB 61 KPKSQAPFRARRTTYAPBPENNALNTOTQPKAHTTDRKEANQAPPEEQDKVPHTAQ 120
QY 728 AAAAGAGGATGCGATTGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRETAQSWKSQDFTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATTAACGTCATCTTGTAGTATGTCACCTTAATGTGACGAGATGCAAAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSHMLAPTGAVSTRTRKQKVTAVIPPKKKPOATPPAPF 240
QY 848 TGGGTTAAATGGCGATGATGATGAGTCCGAAAACAGTTTAACTAGTTAAAGGA 907
DB 241 QSPFTTORNRKLANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSLMLQKFLPN 300
QY 908 AGCTGAAGCTGAACCTGTCAAGATAAGCCTGTGGAATACAGTTGATTAACCAATT 967
DB 301 LTLFLDSHFHMQSEMDRLHFAFPFGFMELNYSLVQKVTFRPPVPOQQLLASLPASL 360
QY 968 CAGTGTCAAGGAGTCTCTGATATGCTGAGCTTCCCTGTGCGATGTGGGTAATTA 1027
DB 361 RCITCAVANGGILNNSHMQEIDSHDYFRLSGLIKGYEQDVGTRTSFYGTAFFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTTGAATCTAACTATTCGGTAGAG 1087
DB 421 SLILGNRGFKNVPLGKQVRYLHLEGTIDYEMLEALLMNOTVMSKNLFWFRRHPOEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLLHPD 494

RESULT 6
US-10-158-790-462
Sequence 462, Application US/10158790
Publication No. US20030180879A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel

;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P330R1C448
;; CURRENT APPLICATION NUMBER: US/10/158,790
;; CURRENT FILING DATE: 2002-05-30
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 462
;; LENGTH: 600
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-158-790-462

Query Match 2.6%; Score 36.4; DB 16; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGGATCTGAGACATTTCTTAAGGCATCCATA 667
DB 1 MRSCLMCRHLISQGVMSLLAVLVFLFALPSPFIKEPQTPSRHQRTEIKERSLSLSLA 60
QY 668 TAGGCATTGGCGCGTAAGTCCGATTTGAAGGATTAAGGGGGGTGTAAGTGTGTCTC 727
DB 61 KPKSQAPFRARRTTYAPBPENNALNTOTQPKAHTTDRKEANQAPPEEQDKVPHTAQ 120
QY 728 AAAAGAGGATGCGATTGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRETAQSWKSQDFTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATTAACGTCATCTTGTAGTATGTCACCTTAATGTGACGAGATGCAAAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSHMLAPTGAVSTRTRKQKVTAVIPPKKKPOATPPAPF 240
QY 848 TGGGTTAAATGGCGATGATGATGAGTCCGAAAACAGTTTAACTAGTTAAAGGA 907
DB 241 QSPFTTORNRKLANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSLMLQKFLPN 300
QY 908 AGCTGAAGCTGAACCTGTCAAGATAAGCCTGTGGAATACAGTTGATTAACCAATT 967
DB 301 LTLFLDSHFHMQSEMDRLHFAFPFGFMELNYSLVQKVTFRPPVPOQQLLASLPASL 360
QY 968 CAGTGTCAAGGAGTCTCTGATATGCTGAGCTTCCCTGTGCGATGTGGGTAATTA 1027
DB 361 RCITCAVANGGILNNSHMQEIDSHDYFRLSGLIKGYEQDVGTRTSFYGTAFFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTTGAATCTAACTATTCGGTAGAG 1087
DB 421 SLILGNRGFKNVPLGKQVRYLHLEGTIDYEMLEALLMNOTVMSKNLFWFRRHPOEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLLHPD 494

RESULT 7
US-10-137-871-462
Sequence 462, Application US/10137871
Publication No. US20030207350A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven

```

; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-462

```

```

Query Match          2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCACATCTTAAGCAGGAGATCTGAGACATTTCTTAAGCATCCATA 667
DB 1 MRSCLMRCHLSQGVMSLLAVLFPLPSPFIKEPQTKPSRHQRTENIKERLSLSLA 60
QY 668 TAGGATTTGGGCGGCTTAAGCGGATTTGAAGAGATTAAGGGGGGTGGAAGTGTGTCTC 727
DB 61 KPKSGAPRARRRTTYABAPENNALNTQTPKXHTTDRKCANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATATACAGCGGCTAAGCAGGAGGTGGCTGACAGCTGTGCAGC 787
DB 121 RAAMWSPEKEXKTMVNTLSRPGDAGMASRTEAOSKSDPTTGTGNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGCTCACTTGTGATGATGTCACCTAATGTGACAGATGCAATGCTGAT 847
DB 181 SEKHGKATTAATKTLIPKSHMLAPTVGAVSTRTRQKQVTTAVIPPEKKPQATPPAPAF 240
QY 848 TGGGTTAAATGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
DB 241 QSPPTQRNORLKANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASKSLMLQKLPFN 300
QY 908 AGCTGAAGCTGAACCTGTCAAGAAATAAGCTGTGGAATACAGTTGATACCAAT 967
DB 301 LTLFLDSRHFNOSEWDRLEHFAAPRPFEMELNYSLVQKVTRPPVPOQOLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGTCTGATATGCTGAGCTTCCCTGCGATTTGGGGTAACTATT 1027
DB 361 RCITCAVANGGILNNSHMGDEIDSHDVYFRISGALIKGYEDVGTRTSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTATTTTCAATGAATCAATCAATCAATTTCTGGTAGAG 1087
DB 421 SLLILGNRGFKNVPKGVRYLHPLFEGTRDYEMLEALLMNQVMSKNLFWFHRHROEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 8
US-10-140-923-462
; Sequence 462, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C18
; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-923-462

```

```

Query Match          2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCACATCTTAAGCAGGAGATCTGAGACATTTCTTAAGCATCCATA 667
DB 1 MRSCLMRCHLSQGVMSLLAVLFPLPSPFIKEPQTKPSRHQRTENIKERLSLSLA 60
QY 668 TAGGATTTGGGCGGCTTAAGCGGATTTGAAGAGATTAAGGGGGGTGGAAGTGTGTCTC 727
DB 61 KPKSGAPRARRRTTYABAPENNALNTQTPKXHTTDRKCANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATATACAGCGGCTAAGCAGGAGGTGGCTGACAGCTGTGCAGC 787
DB 121 RAAMWSPEKEXKTMVNTLSRPGDAGMASRTEAOSKSDPTTGTGNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGCTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
DB 181 SEKHGKATTAATKTLIPKSHMLAPTVGAVSTRTRQKQVTTAVIPPEKKPQATPPAPAF 240
QY 848 TGGGTTAAATGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
DB 241 QSPPTQRNORLKANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASKSLMLQKLPFN 300
QY 908 AGCTGAAGCTGAACCTGTCAAGAAATAAGCTGTGGAATACAGTTGATACCAAT 967
DB 301 LTLFLDSRHFNOSEWDRLEHFAAPRPFEMELNYSLVQKVTRPPVPOQOLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGTCTGATATGCTGAGCTTCCCTGCGATTTGGGGTAACTATT 1027
DB 361 RCITCAVANGGILNNSHMGDEIDSHDVYFRISGALIKGYEDVGTRTSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTATTTTCAATGAATCAATCAATCAATTTCTGGTAGAG 1087
DB 421 SLLILGNRGFKNVPKGVRYLHPLFEGTRDYEMLEALLMNQVMSKNLFWFHRHROEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 9
US-10-141-756-462
; Sequence 462, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```

```
APPLICANT: Gao,Wei-Qiang
APPLICANT: Gerritsen,Mary E.
APPLICANT: Goddard,Audrey
APPLICANT: Godowski,Paul J.
APPLICANT: Gurney,Austin L.
APPLICANT: Sherwood,Steven
APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C200
CURRENT APPLICATION NUMBER: US/10/141,756
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 462
LENGTH: 600
TYPE: PRT
ORGANISM: Homo Sapien
US-10-141-756-462
```

```
Query Match 2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred.No.1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
QY 608 AGACCTACATGATGCGCATCTTAAAGCAGGATCTGAGACATTTCTAAGCATCCATA 667
DB 1 MRSCLMCRHLSQGVMSLLAVLFLPALPSFIKEPQTRSHQRTEINIKERSLSLSIA 60
QY 668 TAGCATTGGGGCGGCGTAAGTCGCGCATTAAGAGAGATAGGGGGGTGTAAGGTGTGTC 727
DB 61 KPSQAPFRRARTTYABEPANNALNTOTQPKAHTTGDREKANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTCGATGGCTATATACAGCCGCTAAGCAGGTGGGCTAAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEASQWSQDRTKTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGTCATCTGTTAGTATGTCACCTAATGTGACGATGCAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPEKKEQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAACAGGATTAAGTATGATTAAGGGA 907
DB 241 QSPITQRNORLKAANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASISLMLQKLFPLN 300
QY 908 AGCTGAAGCTGAACCTGTCAGAATAAAGCTGTTGAATACAGCTTGATTAACCAATT 967
DB 301 LTLFLDSHRHFNQSEMDRLHFAFPFGFMELNYSLVQKVTRFPVPVQQQLLASLPAAGSL 360
QY 968 CAGTGTCAAGGGTGTCTGATATGCTGAGCTTCCCTGTCCATTTGGGGTAATCTATT 1027
DB 361 RCITCAVVGNGILNNSHMGDEIDSHDVYFRLSGALIKGYBODVGTSTRSFYFTAFSLQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTAATGAATCTAAGTATCTGGGTAGAG 1087
DB 421 SLLILGNRFKNVPLGKDVRYLHLEGRDYEMLEALMNQTVMSKNLFWFRHRQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494
```

```
RESULT 10
US-10-141-759-462
; Sequence 462, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

```
APPLICANT: DeForge,Laura
APPLICANT: Desnoyers,Luc
APPLICANT: Filvaroff,Ellen
APPLICANT: Gao,Wei-Qiang
APPLICANT: Gerritsen,Mary E.
APPLICANT: Goddard,Audrey
APPLICANT: Godowski,Paul J.
APPLICANT: Gurney,Austin L.
APPLICANT: Sherwood,Steven
APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C197
CURRENT APPLICATION NUMBER: US/10/141,759
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 462
LENGTH: 600
TYPE: PRT
ORGANISM: Homo Sapien
US-10-141-759-462
```

```
Query Match 2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred.No.1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
QY 608 AGACCTACATGATGCGCATCTTAAAGCAGGATCTGAGACATTTCTAAGCATCCATA 667
DB 1 MRSCLMCRHLSQGVMSLLAVLFLPALPSFIKEPQTRSHQRTEINIKERSLSLSIA 60
QY 668 TAGCATTGGGGCGGCGTAAGTCGCGCATTAAGAGAGATAGGGGGGTGTAAGGTGTGTC 727
DB 61 KPSQAPFRRARTTYABEPANNALNTOTQPKAHTTGDREKANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTCGATGGCTATATACAGCCGCTAAGCAGGTGGGCTAAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEASQWSQDRTKTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGTCATCTGTTAGTATGTCACCTAATGTGACGATGCAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPEKKEQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAACAGCTTGAATCTAATTAAGGGA 907
DB 241 QSPITQRNORLKAANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASISLMLQKLFPLN 300
QY 908 AGCTGAAGCTGAACCTGTCAGAATAAAGCTGTTGAATACAGCTTGATTAACCAATT 967
DB 301 LTLFLDSHRHFNQSEMDRLHFAFPFGFMELNYSLVQKVTRFPVPVQQQLLASLPAAGSL 360
QY 968 CAGTGTCAAGGGTGTCTGATATGCTGAGCTTCCCTGTCCATTTGGGGTAATCTATT 1027
DB 361 RCITCAVVGNGILNNSHMGDEIDSHDVYFRLSGALIKGYBODVGTSTRSFYFTAFSLQ 420
QY 1028 TCATAGTGGGCGAGATGCAACTCTATTTCATTAATGAATCTAAGTATCTGGGTAGAG 1087
DB 421 SLLILGNRFKNVPLGKDVRYLHLEGRDYEMLEALMNQTVMSKNLFWFRHRQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494
```

```
RESULT 11
US-10-140-805-462
; Sequence 462, Application US/10140805
; Publication No. US20030207417A1
```



```
RESULT 13
US-10-142-426-462
; Sequence 462, Application US/10142426
; Publication No. US20040048333A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P333081C224
; CURRENT APPLICATION NUMBER: US/10/142,426
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-426-462

Query Match
Best Local Similarity 2.6%; Score 36.4; DB 17; Length 600;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AACCTACATGATGCCACATCTTAAAGCAGGATCGACATTTTCTAAGCATCCANA 667
DB 1 MSCSMRCHLSQGVWSLLAVLFPLALPSFIKEPQTKPSRIQTEINERISLOSTIA 60
QY 668 TAGGATTTGGGCGCTAATGCGCATTTGAAGAGATGAAGGGGCTGAAGTGTGTCTC 727
DB 61 KKSQAIPRAKRTTYIABEPAPENMALNTOTOPKANTTGRKEANQAPREBODKVPHTQ 120
QY 728 AAAAGAGGTGATTTGGCTATACCAAGCGCTAAGCAGGTGGGCTAGACAGCTGTGCACG 787
DB 121 RAAMKSPKEXTMVTLLSPRGDAGMASGRTEAQSWKSDTKTQONGGQTKRLTFASRTV 180
QY 788 TGTGAATACGTCACCTGCTTAGGATGTGCTCACCTTAATGTCAGCAGATCAATGTGTAT 847
DB 181 SEKHOGKATTTKTLIPKSOHRLMALTGTAVSTRTRQKGVTVIIPKEKKPQATPPAPF 240
QY 848 TGGGTTAAATGGGAGTGAAGTAGTAGTGGCCGAAAAACAGCTTAACTAGTTAAAGGGA 907
DB 241 QSPPTORNRKLANFKSPRWDPEEKYSFEIGLQTTTPDSVKIKASSLMLQKLFLEN 300
QY 908 AGCTGAAAGCTGAACCTGTCAAGAAATTAAGCTGTGGAATAAGCACTTGATACCAANT 967
DB 301 LTLPLDSRRHFNQSEMDRLHFAPRPFMELNYSLVQKVTTRPPVQQQLLASLPAGSL 360
QY 968 CAGTGTCTCAGGCTGCTGATATGCTGAGCTTCCCTGTGCGATTTGGGGTAACTATT 1027
DB 361 RCITCAVVGNGGILNNSHMGDEIDSHDYFRLGALINKYEDVDGTRTSFYFTAFSLQ 420
QY 1028 TCATAGTGGGAGATGAACCTATTTTCAATTGAATCACTTAACATTTCTGGTAGAG 1087
DB 421 SLILGNRGPKNVPJGKDVRYLHPLEGTIDYEWLEALLMNQTVMSKNLFWFHRPOEAFR 480
```

```
QY 1088 TTCTCATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 14
US-09-775-938A-34/C
; Sequence 34, Application US/09775938A
; Patent No. US20020081665A1
; GENERAL INFORMATION:
; APPLICANT: Haygood, M.
; APPLICANT: Davidson, S.K.
; APPLICANT: Allen, S.W.
; APPLICANT: Hildebrand, M.
; TITLE OF INVENTION: Bryostatins, Bryopyrans and Polyketides: Compositions and Methods
; FILE REFERENCE: 1133.010US1
; CURRENT APPLICATION NUMBER: US/09/775,938A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: PCT/US00/21326
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/147,283
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2672
; TYPE: DNA
; ORGANISM: Endobugula sertula
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2672)
; OTHER INFORMATION: N refers to any nucleotide.
US-09-775-938A-34

Query Match
Best Local Similarity 2.6%; Score 36.2; DB 9; Length 2672;
Matches 41; Conservative 60; Mismatches 69; Indels 0; Gaps 0;

QY 674 TTGGCGGTGAATGCGCATTTAAGAGATTAAGGGGCTGTGAAGAGTGTGTCAAAAG 733
DB 702 TTCMGCMGRSTAKSMYKTKYKWTWCCAGGSTRKSKGTWWTYGGMANCCRG 643
QY 734 AGGTGATTTGGCTATACCAAGCCGCTAAGCAGTGGGCTAGACAGCTGTGACGCTGAA 793
DB 642 MYSWGTTTTCRAAWKSSGYTKCGTGTAGTGWCCAWWMSCRMWSCRSGSGKCTTKWT 583
QY 794 TAACGCTAGTGTATGTCACCTAATGTCAGCAGATGCAAAATGC 843
DB 582 KAAWAMCTYTTTGSAAWAASTYTKWAGGWKMGCGKGMCRAMTKY 533

RESULT 15
US-10-312-841-1/C
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomic AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 367378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
```


Query Match 2.6%; Score 36; DB 16; Length 3673778;
 Best Local Similarity 56.9%; Pred. No. 3.3e+02;
 Matches 66; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 1243 AAAAGAGCTTGAGAGACCATTCAAAATCCTAAGGCTCTCTTTCTTTCTGATCAATCA 1302
 Db 604282 AAAATTTACATTAATTAATCATTTCTAATCATATCTCCAAACTTATATCTTTCCCGCTA 604223
 QY 1303 AGAATCATACACTCAACGAGACTCTTTCTATCTTCCTATAGCAATTCCCAA 1358
 Db 604222 ACAATTAATACGTTACAACTCTTAACCTATCTACATATATATACAAATATCTTA 604167

Search completed: February 9, 2005, 15:26:37
 Job time : 769 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:10:25 ; Search time 4286 Seconds
(without alignments)
12238.116 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtgcagctg99t9gtgatac.....accatcaatcaactaaca 1378

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	3.0	557	8	BH285801 CH230-118
2	39	2.8	1031	9	CNS0500M
3	38.8	2.8	1205	8	CC264185 CH261-57E
4	38.4	2.8	673	8	BZ246183 CH230-374
5	38.4	2.8	734	9	AG590975 Mus muscu
6	38.2	2.8	479	8	AO018895 CIT-HSP-2
7	38.2	2.8	923	8	CNS0404PT
8	38	2.8	482	4	BG185531
9	37.8	2.7	1101	9	CNS0000D1
10	37.6	2.7	688	1	AV384223
11	37.4	2.7	598	8	AZ960759
12	37.2	2.7	250	9	CE251889
13	37.2	2.7	625	7	CN084563
14	37	2.7	446	9	CL700776
15	37	2.7	590	9	CE571358
16	37	2.7	725	9	CL705705 SP_Bb004
17	37	2.7	620	9	CE686685
18	37	2.7	801	4	BG534705
19	36.8	2.7	562	9	CE834892
20	36.8	2.7	704	9	CL197184
21	36.8	2.7	730	7	CN975554
22	36.8	2.7	738	9	CL157592
23	36.8	2.7	773	9	CL157591
24	36.8	2.7	1101	9	CNS016K9

C 25	36.6	2.7	1075	8	CC280186	CC280186 CH261-112
26	36.4	2.6	778	7	CV490143	CV490143 AGENCOURT
27	36.4	2.6	1030	5	BK460908	BK460908 BX460908
28	36.2	2.6	457	4	BT741102	BT741102 GC93511.Y
29	36.2	2.6	634	8	AZ037814	AZ037814 RPT-23-2
C 30	36	2.6	532	1	AL923574	AL923574 AL923574
31	36	2.6	551	8	AZ163313	AZ163313 SP_0073.B
32	36	2.6	621	5	BP504692	BP504692 BP504692
33	36	2.6	683	1	AA556936	AA556936 778 Lobi
C 34	36	2.6	1201	9	CNS0102D	AA098431 Drosophila
C 35	35.8	2.6	432	4	BT502064	BT502064 tm08105.Y
C 36	35.8	2.6	474	4	BT501745	BT501745 tm04605.Y
37	35.8	2.6	592	8	AZ838544	AZ838544 2M0134H15
38	35.8	2.6	610	8	AO891964	AO891964 HS_3142.A
C 39	35.8	2.6	845	8	BZ114424	BZ114424 CH230-510
40	35.8	2.6	891	6	CD789126	CD789126 EST660487
41	35.8	2.6	945	6	CD792880	CD792880 EST664241
C 42	35.8	2.6	1306	8	BZ565154	BZ565154 PAC82-164
C 43	35.6	2.6	644	8	BH766160	BH766160 BMBAC359H
C 44	35.6	2.6	748	4	BI886660	BI886660 ZF637-1.0
C 45	35.6	2.6	814	8	BH109031	BH109031 RPT-24-3

ALIGNMENTS

RESULT 1
BH285801
LOCUS BH285801 557 bp DNA linear GSS 30-NOV-2001
DEFINITION CH230-118D10.TVB CHORI-230 Segment 1 Rattus norvegicus genomic
clone CH230-118D10, genomic survey sequence.
BH285801
ACCESSION BH285801 GI:17198207
VERSION
KEYWORDS
SOURCE GSS.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 557)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsengaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,L., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EORI segment
Unpublished (1999)
COMMENT Other_GSSs: CH230-118D10.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pierre de Jong (pje@genemail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or erting_information.html). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Place: 118 row: D column: 10
Seq primer: 17
Class: BAC ends.

FEATURES
Location/Qualifiers
1..557
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SaNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-118D10"
/sex="female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;

ORIGIN CHORI-230 Rat (BN/SHNhd/MCM) BAC library produced by
Pleeter de Jong"

Query Match 3.0%; Score 41.6; DB 8; Length 557;
Best Local Similarity 49.1%; Pred. No. 0.66; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
OY 1155 AGAGTGGCGCATTTAGAGTACGACATCGAAGCCGGAAGCGTTCAAGACATGGCGT 1214
DB 89 AGATGCTGGATTAATGAAGATGACAAAGAGAGGGGAGACATATGAAGAGTCAATCAGG 148
OY 1215 ACGTGATACATAGAGTCAATGAAACATTAAGAGCGTTGAAGAACCATTTAAATCTTA 1274
DB 149 ACATGTATGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 208
OY 1275 GGGTCT 1334
DB 209 CGTTATCATGTTTCTTTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 268
OY 1335 ATCTTCCCTATAGCAATTCCTCAAAACCATCAATCACTTACA 1378
DB 269 GACTGACTAGAGAAATCCCAACCTCTAGATACATAGCA 312

RESULT 2
CNS05QOM/c 1031 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence Sp6 end of clone
DEFINITION 025B03 of library B from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL349601 GI:8243371
VERSION GSS; genome survey sequence.
KEYWORDS Tetradon nigroviridis
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradon.
REFERENCE 1 Roest Croollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Winkler, P., Broctier, P., Quetier, F.,
Saurin, W., Bernot, A., Bouneau, L., Billault, A., Quetier, F.,
Fizames, C., Fischer, C., and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater puterfish Tetradon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645
REFERENCE 2
JOURNAL Roest Croollius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A., Bouneau, L., and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater puterfish Tetradon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143
REFERENCE 3
JOURNAL 3 (bases 1 to 1031)
MEDLINE Genoscope.
PUBMED Direct Submission
REFERENCE 4
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read, and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
FEATURES
source location/Qualifiers
1..1031
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"

ORIGIN /clone="025B03"
/clone_lib="B"
/note="Genoscope sequence ID : COAB025CA02B1-end : Sp6"

Query Match 2.8%; Score 39; DB 9; Length 1031;
Best Local Similarity 43.1%; Pred. No. 5.3; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
OY 792 AATAAGCTCACTGCTTATGATGTCACCTTAATGTCAGCAGATGCAATGCTGATTTGG 851
DB 523 AAAAAAGTTGTTCTGCTAGACAGCTNACTAATAATTAACNAGAAANAGNAATAAGT 464
OY 852 TTAATAAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
DB 463 MTNAATGAGNANN 404
OY 912 GAAAGCTGACCTGTCAGAAATTAAGCCGCTGGAATACAGGTTGATACCAATTCAGT 971
DB 403 NAATNCTTACCATTTNAAAAACNGNANNTAATGNNAACTNAGTTCAAAATCCNACTCTTT 344
OY 972 CGTCAAG 979
DB 343 CTTNAGG 336

RESULT 3
CC264185 1205 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-57E7_Sp6.1 CH261 Gallus gallus genomic clone CH261-57E7,
DEFINITION genomic survey sequence.
ACCESSION CC264185
VERSION CC264185.1 GI:30610833
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
REFERENCE 1 (bases 1 to 1205)
Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
JOURNAL Contact: Richard K. Wilson
COMMENT Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 50
High quality sequence stop: 729.
FEATURES
source location/Qualifiers
1..1205
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-57E7"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match 2.8%; Score 38.8; DB 8; Length 1205;
Best Local Similarity 48.2%; Pred. No. 6.4; Indels 0; Gaps 0;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
OY 823 AATGTACAGCATGCAATGCTGATTTGGTTTAAATGGGCAATGATGATGATGATGATGATGAT 882

```

Db      845  AATCCCACTGACAGATTGATTATTTGGGATGAAGGGAAGGCTTGGCGGCGGACGAA 904
Qy      883  AACCCTTGAATCTAGTTAAGGGAAGCTGAAGCTGAACCTGTCAAGAAATAAGCCTGTT 942
Db      905  AAAAGTTAACCCCATTCATTAATGAAGAACCTTCAGTAAACCTTAAGATGACCAACAGA 964
Qy      943  GGAATACACAGTTGATTAACCAATTCAGTCAAGAGGTGCTGTATATGCTGAAGCTTC 1002
Db      965  AAAAATAAAGGGGGTCTATCAACCTTGTTATGAAGGGATCCCATTAATAGTTTCCTTA 1024
Qy      1003 CCTGTGCATTGTGGGGTAACTATTTCATAGTGGGCGACAAATGCAA 1048
Db      1025 ACCCTCATTCGCGGCAACAAAATTCAATGACGGGCGAAGAAATAA 1070

RESULT 4
BZ246183/c  BZ246183  673 bp  DNA  linear  GSS 12-OCT-2002
LOCUS      CH230-374A18.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-374A18, genomic survey sequence.
ACCESSION  BZ246183
VERSION    BZ246183.1 GI:23906447
KEYWORDS   GSS.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 673)
REFERENCE  Zhao,S., Shetty,U., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.,
            Riggs F., de Jong P. and Fraser,C.M.,
            Rat BAC End Sequences from Library CHORI-230 MboI segment
            Unpublished (1999)
            Other GSSs: CH230-374A18.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@igr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering information.html). BAC end
            page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
            Plate: 374 row: A column: 18
            Seg primer: T7
            Clones: BAC ends.
FEATURES
    source          Location/Qualifiers
        1..673
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /strain="BN/SsNHsd/MCw"
            /db_xref="taxon:10116"
            /clone="CH230-374A18"
            /sex="Female"
            /cell_type="Brain"
            /clone_lib="CHORI-230 Segment 2"
            /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
            CHORI-230 Rat (BN/SsNHsd/MCw) BAC library produced by
            Pieter de Jong"
ORIGIN
Query Match      2.8%; Score 38.4; DB 8; Length 673;
Best Local Similarity 51.1%; Pred. No. 6.9;
Matches 90; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Qy      1100 TTCTGCTGTCACTTACACATCATGCGGGTCAACACGTTATACAGTTCTTAGAGACT 1159
|||||

```

```

Db      606  TTCTGTTTTAAATTAATGAATAATGCTATTATGATATATATATATACACACAGA 547
Qy      1160 GCGGCAATTAAGTACTACCGCATGAACCCGGAAGCGGTTCAGACATGCGCTACGTA 1219
Db      546  GAGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 487
Qy      1220 GATACATAGACTCATAGAAACATTAAGAGAGCTTGAAAGACATTCMAATCCTAAG 1275
Db      486  CACAGATACACACACACACACACAGAGAGAGATTCTGCMAAGCTTCAAGAGAAAG 431

RESULT 5
AGS90975/c  AGS90975  734 bp  DNA  linear  GSS 05-JUN-2004
LOCUS      Mus musculus molossinus DNA, clone: MSMG01-519M14.T, genomic survey
DEFINITION sequence.
ACCESSION  AGS90975
VERSION    AGS90975.1 GI:48351805
KEYWORDS   GSS.
SOURCE     Mus musculus molossinus
            Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1
REFERENCE  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
            BAC end Sequences of library MSMG01
            Unpublished
            2 (bases 1 to 734)
REFERENCE  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
            Direct Submision
            Submitted (17-NOV-2003) Masehiya Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suhiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel:81-45-503-9111, Fax:81-45-503-9170
            E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170
            Clones are derived from the mouse BAC library MSMG01. For BAC
            library availability, please contact Kunya Abe (abe@rtc.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector : pBACe3.6
            R.Site 1 : EcoRI
            R.Site 2 : EcoRI.
FEATURES
    source          Location/Qualifiers
        1..734
            /organism="Mus musculus molossinus"
            /mol_type="genomic DNA"
            /sub_species="molossinus"
            /db_xref="taxon:57486"
            /clone="MSMG01-519M14.T7"
            /sex="male"
            /tissue_type="mixture of kidney and spleen"
            /clone_lib="MSMG01 Mouse Male BAC Library"
ORIGIN
Query Match      2.8%; Score 38.4; DB 9; Length 734;
Best Local Similarity 54.2%; Pred. No. 7.1;
Matches 78; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
Qy      1106 CTGTCACTTACACACATCATGGGGTCAACACGTTATACAGTTATAGAGTGGCGCA 1165
Db      670  CAGACAGACACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 611
Qy      1166 TTGAAGTACTATCCGATTCGAACCCGGAAGCGTTCAAGACATGCGCGTACTAGATCA 1225
Db      610  GACACACAGATTAAGAGACACACAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGACA 551
Qy      1226 TAGAGTCATAGAAACATAAAGGA 1249
|||||

```

Db 550 GAGAGATACAGTACGACAGAGA 527

RESULT 6
LOCUS A0018895/c 479 bp DNA 1linear GSS 09-JUN-1998
DEFINITION CIT-HSP-2302116.TF CIT-HSP Homo sapiens genomic clone 2302116,
genomic survey sequence.
ACCESSION A0018895
VERSION A0018895.1 GI:3197631
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 479)
Adams,M.D., Rounley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
JOURNAL Other_GSSes: CIT-HSP-2302116.TR
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Classes: BAC ends.

FEATURES
source
1..479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2302116"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match 2.8%; Score 38.2; DB 8; Length 479;
Best Local Similarity 50.3%; Pred. No. 7;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 31 GGTGTCGAAGCTGTGTAGTACGATGAGAAAGAAAGCGTGTGGCAGCT 90
DB 460 GGTGTCGAAGCTGTGTAGTACGATGAGAAAGAAAGAAAGCGTGTGGCAGG 401
QY 91 GAAGCTGAAGAGAGAGCAAGATTAATGACATGCGATGCGTTCATGCTTGT 150
DB 400 AGTGAAGAAAGTAGAGGGTATGATATATGAAATGCTTATATCGCTGTGGC 341
QY 151 TCAAGAGACAGCTTGATCTACCTGGTGTCCCTCTTCGTTGTAACAAGATCAATATC 210
DB 340 TAGGCAATGCACTTGTATGATTCCTCTCTCTTTGTTGAAATGTTCAATCAT 281
QY 211 GGATGAC 217
DB 280 GGAAGAA 274

RESULT 7
LOCUS CNS04NPT/c 923 bp DNA 1linear GSS 01-SEP-2000

DEFINITION Tetradon nigroviridis genome survey sequence PUC-ori end of clone
123C09 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL299018
VERSION AL299018.1 GI:8038159
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,M. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
2
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,M., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
JOURNAL MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 923)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
Location/Qualifiers
source
1..923
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="123C09"
/clone_lib="G"
/note="Genoscope sequence ID : COB6123AB05SP1-end :
PUC-ori"

ORIGIN
Query Match 2.8%; Score 38.2; DB 9; Length 923;
Best Local Similarity 50.3%; Pred. No. 8; 9;
Matches 85; Conservative 3; Mismatches 81; Indels 0; Gaps 0;

QY 891 AGATCTAGTTAAAGGAAGCTGAAGCTGAACCTGTCAAGAAATTAAGCTGTGAAATCA 950
DB 352 RGAGAAAGAGAAAAAAGCCATGCTGTGTATACAAACATATGTGTGGCTCG 293
QY 951 ACGTGAATACCCCAATTCAGTCGTCAAGGGTGTCTGTATATGCTGAGACTTCCCTGTGGC 1010
DB 292 GCCTTAATATCTTCAAAATGTTGGATGAAAGCCGACTGGCGGCTTTCGCTTAAT 233
QY 1011 ATTGGGGGTAACTATTTGATAGTGGGAGAAATGCACTGATTTTCA 1059
DB 232 ATCTCCGCTCCCTGTGATTGCGCGGTGAACAAAGACAAATATCA 184

RESULT 8
LOCUS BG185531/c 482 bp mRNA 1linear EST 21-APR-2001
DEFINITION RST4479 Athysys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG185531

VERSION	EST.	GI:13707218
KEYWORDS	BJ185531.1	GI:13707218
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 482)	
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McMillgott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Kilka,A., Hees,D., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.	
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression	
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)	
MEDLINE	21227151	
PUBMED	11329013	
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scain@athersys.com High quality sequence stop: 451. Location/Qualifiers 1..482	
FEATURES	source /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="HT1080" /clone_lib="Athersys RAGE library" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology', in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."	
ORIGIN	Query Match Best Local Similarity 2.8%; Score 38; DB 4; Length 482; Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;	
QY	913 AAGCTGAAACCTGTCAGAAATAGCCTTGGAATACACCTGTGATACCAATTCAGTC 972	
Db	164 AGAGCAACCCCTTTCTGTACAGACATTTGAANTCATCCATGCTCATTTGGAT 105	
QY	973 GTCAAGGCTGCTGATATGCTGAGCTTCCTGTGCATTGGGGTAACATATTCATA 1032	
Db	104 TTCAAGTGTTCCTCCCTCCCACTCNNAACGCGATTCGGAGTTTGGAAATGTTGGAA 45	
QY	1033 GTGGGGCA 1040	
Db	44 GTATTGNA 37	
RESULT 9		
CNS000D1/c		
LOCUS	CNS000D1	1101 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR01J16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL065414	
VERSION	AL065414.1	GI:4938827
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 1101)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	

[illegible]

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

```
/moi_type="genomic DNA"  
/strain="Standard Poodle"  
/db_xref="taxon:9615"  
/clone_lib="Dog Library"
```


QY 329 AACGAAATACACGTTGTA 349
 |||||
 Db 126 AAAGAGAAGCATTAAGAGAGAA 146

RESULT 15

CE571358 590 bp DNA linear GSS 28-SEP-2003
 LOCUS clgr-gss-dog-17000327517076 Dog Library Canis familiaris genomic,
 DEFINITION genomic survey sequence.

CE571358
 ACCESSION CE571358.1 GI:36888139
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Canis familiaris (dog)
 Canis familiaris
 Canis familiaris

REFERENCE
 AUTHORS
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 590)
 Kirkness,E.F., Balna,V., Halpern,A.L., Levy,S., Remington,K.,
 Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
 Venter,J.C.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 The dog genome: survey sequencing and comparative analysis
 Science 301 (5641), 1898-1903 (2003)
 22873432
 14512627

COMMENT
 Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkness@ligr.org
 Class: shotgun.

FEATURES
 source
 Location/Qualifiers

1..590
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site 1: BsrXI; Libraries were prepared from
 peripheral blood"

ORIGIN

Query Match 2.7%; Score 37; DB 9; Length 590;

Best Local Similarity 58.7%; Pred. No. 18;
 Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1218 TAGATACATAGAGTATAGAAACATAAAGAGCTGAAGACCATTCAAATCCTAAGGG 1277
 |||||
 Db 326 TACCTAAATAGAACATATAATATATATGAGATTCAAAAAATTAATTACTTCCCAAGGG 267
 |||||
 QY 1278 TCTCTCTTTCTTCGATCAGATCAAGATCATACATCAACGAGAA 1326
 |||||
 Db 266 GCTTTTGCATCTGATCTCATCTGTATATTTCAGAAAGATCAAAAA 218
 |||||

Search completed: February 9, 2005, 13:39:56
 Job time : 4291 secs